

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:19:43 ; Search time 27.08 Seconds
(without alignments)
28.387 Million cell updates/sec

Title: 10-071247-1

Perfect score: 59

Sequence: 1 gywgkyw 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	79.7	1365	1 BVBYK5	killer toxin resis
2	45	76.3	441	2 F83694	short-chain fatty
3	43.5	73.7	83	2 E69903	hypothetical prote
4	43.5	73.7	187	2 G83047	hypothetical prote
5	43	72.9	358	2 H83554	hypothetical prote
6	43	72.9	470	2 B72867	hypothetical prote
7	42	71.2	80	2 AC2394	hypothetical prote
8	42	71.2	180	2 D83812	hypothetical prote
9	42	71.2	181	2 A83915	hypothetical prote
10	42	71.2	196	2 AE2918	acetyltransferase
11	42	71.2	207	2 H97692	conserved hypothet
12	42	71.2	563	2 A69491	probable formylmet
13	42	71.2	1311	2 A56390	mannosyl-glycoprot
14	42	71.2	1312	2 E95006	beta-N-acetylhexos
15	42	71.2	1312	2 A97879	beta-N-acetylhexos
16	41	69.5	113	2 PH1663	lg heavy chain V r
17	41	69.5	145	2 F75337	transposase - Dein
18	41	69.5	375	2 A12041	hypothetical prote
19	41	69.5	844	2 I37079	DNA ligase (ATP) (
20	41	69.5	1447	2 S63669	UDPglucose-glycop
21	40	67.8	489	2 S77357	hypothetical prote
22	40	67.8	495	2 AD1927	hypothetical prote
23	40	67.8	554	1 A28716	saposin precursor
24	40	67.8	589	2 B29514	muscarinic acetylch
25	40	67.8	589	2 A29476	muscarinic acetylch
26	40	67.8	590	2 S01114	muscarinic acetylch
27	40	67.8	590	2 S47572	muscarinic acetylch
28	40	67.8	741	2 S73827	hypothetical prote
29	39	66.1	102	2 B72742	hypothetical prote

ALIGNMENTS

RESULT 1

BVBYK5

killer toxin resistance protein KRE5 precursor - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein Q6254; protein YOR336W
C:Species: Saccharomyces cerevisiae
C:Date: 30-Sep-1991 #sequence_revision 10-May-1996 #text_change 16-Jun-2000
C:Accession: S62066; A36327; S71974; S12202
R:Parle, A.G.; Hand, N.J.; Goulding, S.G.; Wolfe, K.H.
submitted to the EMBL Data Library, June 1995
A:Description: Sequence of 29 kilobases around the PDR10 locus on the right arm of Sa
A:Reference number: S62058
A:Accession: S62066
A:Molecule type: DNA

A:Residues: 1-1365 <PAR>

A:Cross-references: EMBL:Z49821; NID:g1163062; PIDN:CAA89981.1; PID:g1163070

R:Meaden, P.; Hill, K.; Wagner, J.; Slipetz, D.; Sommer, S.S.; Bussey, H.

Mol. Cell. Biol. 10, 3013-3019, 1990

A:Title: The yeast KRE5 gene encodes a probable endoplasmic reticulum protein require
A:Reference number: A36327; MUID:90258892
A:Accession: A36327

A:Molecule type: DNA

A:Residues: 1-581,583-779, 'I', 781, 'IKMKCKONISK', 794, 'K', 795-1365 <MEA>

A:Cross-references: EMBL:M33556; NID:g171794; PIDN:AAA34725.1; PID:g171795

R:Goulding, S.E.; Hand, N.J.; Parle-McDermott, A.G.; Wolfe, K.H.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67233

A:Accession: S67243

A:Molecule type: DNA

A:Residues: 1-1365 <GOU>

A:Cross-references: EMBL:Z75244; NID:g1420730; PIDN:CAA99659.1; PID:g1420731; GSPDB:G

A:Experimental source: strain S288C

R:Parle-McDermott, A.G.; Hand, N.J.; Goulding, S.E.; Wolfe, K.H.

Yeast 12, 999-1004, 1996

A:Title: Sequence of 29 kb around the PDR10 locus on the right arm of Saccharomyces c

A:Reference number: S71966; MUID:97051586

A:Accession: S71974

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1365 <PAW>

A:Cross-references: EMBL:Z49821; NID:g1163062; PIDN:CAA89981.1; PID:g1163070

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995

C:Genetics:

A:Gene: SGD:KRE5; MIPS:YOR336W

A:Cross-references: SGD:S0005863; MIPS:YOR336W

A:Map position: 15R

C:Function:

A:Description: required for normal cell growth

A:Pathway: (1->6)-beta-D-glucan biosynthesis

C:Superfamily: KRE5 protein

C:Keywords: endoplasmic reticulum; glycoprotein

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-1365/Product: killer toxin resistance protein KRE5 #status predicted <MAT>

acetyltransferase,
conserved hypothet
acetyltransferase,
hypothetical prote
hypothetical prote
secretory protein-
probable heme tran
probable phosphoe
hypothetical prote
hypothetical prote
INDAL protein - fu
cyclomaltodextrin
probable membrane
G protein-coupled
probable ABC trans
protein F18014.29

30 39 66.1 177 2 D87288
31 39 66.1 185 2 B97933
32 39 66.1 186 2 A95066
33 39 66.1 194 2 H28445
34 39 66.1 228 2 S26009
35 39 66.1 247 2 T47883
36 39 66.1 250 1 S70854
37 39 66.1 270 1 G83242
38 39 66.1 396 2 T04561
39 39 66.1 499 2 T36462
40 39 66.1 573 1 S33212
41 39 66.1 653 1 ALKBG
42 39 66.1 961 2 S67568
43 39 66.1 962 2 JC5808
44 38.5 65.3 273 2 E95268
45 38 64.4 115 2 B86328

F:1362-1365/Region: endoplasmic reticulum retention signal #status predicted F:115,228,293,457,519,523,604,644,870,1091,1150,1195/Binding site: carbohydrate (Asn) (C)		
Query Match	79.7%; Score 47; DB 1; Length 1365;	
Best Local Similarity	75.0%; Pred. No. 31;	
Matches	6; Conservative 1; Mismatches 0; Indels 1; Gaps 0;	
Qy	1 gywgkgyw 8 	
Db	1234 GYWGEGYW 1241	
RESULT	2	
F83694	short-chain fatty acids transporter atoE [imported] - Bacillus halodurans (strain C-125)	
C:Species:	Bacillus halodurans	
C:Date:	01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001	
C:Accession:	F83694	
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira		
Nucleic Acids Res.	28, 4317-4331, 2000	
A:Title:	Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and	
A:Reference number:	A83650; MUID:20512582; PMID:11058132	
A:Accession:	F83694	
A:Status:	preliminary	
A:Molecule type:	DNA	
A:Residues:	1-441 <STO>	
A:Cross-references:	GB:AP001508; GB:BA000004; NID:g10172890; PIDN:BAB04077.1; GSPDB:GN00	
A:Experimental source:	strain C-125	
C:Genetics:		
A:Gene:	atoE	
C:Superfamily:	conserved hypothetical integral membrane protein HP0693	
Query Match	76.3%; Score 45; DB 2; Length 441;	
Best Local Similarity	71.4%; Pred. No. 22;	
Matches	5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;	
Qy	2 ywgkgyw 8 	
Db	48 YWGEFW 54	
RESULT	3	
E69903	hypothetical protein yodI - Bacillus subtilis	
C:Species:	Bacillus subtilis	
C:Date:	05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999	
C:Accession:	E69903	
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte		
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch		
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.		
Nature	390, 249-256, 1997	
A:Authors:	Foulger, D.; Fritz, C.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen	
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.		
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinols		
A:Authors:	Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel	
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle		
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon		
A:Authors:	Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron	
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,		
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K		
A:Authors:	Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Dauchin, A.	
A:Title:	The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.	
A:Reference number:	A69580; MUID:98044033	
A:Status:	preliminary; nucleic acid sequence not shown; translation not shown	
A:Molecule type:	DNA	
A:Residues:	1-83 <KUN>	
A:Cross-references:	GB:Z99114; GB:AL009126; NID:g2634230; PIDN:CAB13852.1; PID:el185432;	
A:Experimental source:	strain 168	
C:Genetics:		
A:Gene:	yodI	
Query Match	73.7%; Score 43.5; DB 2; Length 83;	
Best Local Similarity	87.5%; Pred. No. 8.5;	
Matches	7; Conservative 0; Mismatches 0; Indels 1; Gaps 1;	
Qy	1 gywgkgyw 8 	
Db	53 GYWG-GYW 59	
RESULT	4	
G83047	hypothetical protein PA4793 [imported] - Pseudomonas aeruginosa (strain PA01)	
C:Species:	Pseudomonas aeruginosa	
C:Date:	15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000	
C:Accession:	G83047	
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;		
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L		
Nature	406, 959-964, 2000	
A:Title:	Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa	
A:Reference number:	A82950; MUID:20437337	
A:Accession:	G83047	
A:Status:	preliminary	
A:Molecule type:	DNA	
A:Residues:	1-187 <STO>	
A:Cross-references:	GB:AE004892; GB:AE004091; NID:g9951049; PIDN:AAG08179.1; GSPDB:GN	
A:Experimental source:	strain PA01	
C:Genetics:		
A:Gene:	PA4793	
Query Match	73.7%; Score 43.5; DB 2; Length 187;	
Best Local Similarity	87.5%; Pred. No. 17;	
Matches	7; Conservative 0; Mismatches 0; Indels 1; Gaps 1;	
Qy	1 gywgkgyw 8 	
Db	109 GYWG-GYW 115	
RESULT	5	
H83554	hypothetical protein PA0736 [imported] - Pseudomonas aeruginosa (strain PA01)	
C:Species:	Pseudomonas aeruginosa	
C:Date:	15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000	
C:Accession:	H83554	
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;		
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L		
Nature	406, 959-964, 2000	
A:Title:	Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa	
A:Reference number:	A82950; MUID:20437337	
A:Accession:	H83554	
A:Status:	preliminary	
A:Molecule type:	DNA	
A:Residues:	1-358 <STO>	
A:Cross-references:	GB:AE004508; GB:AE004091; NID:g9946611; PIDN:AAG04125.1; GSPDB:GN	
A:Experimental source:	strain PA01	
C:Genetics:		
A:Gene:	PA0736	
Query Match	72.9%; Score 43; DB 2; Length 358;	
Best Local Similarity	75.0%; Pred. No. 36;	
Matches	6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
Qy	1 gywgkgyw 8 	
Db	198 GTWGGYW 205	

RESULT 6

hypothetical protein APE1820 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: B72567
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahara, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kawarabayashi, Y. 1999
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain K1
A:Reference number: A72450; MUID:99310339
A:Accession: B72567
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-470 <RAW>
A:Cross-references: DDBJ:AF000062; NID:g5105244; PIDN:BAA80823.1; PID:g5105510
A:Experimental source: strain K1
C:Genetics:
C:Gene: APE1820
C:Superfamily: phytoene dehydrogenase

Query Match	72.9%;	Score 43;	DB 2;	Length 470;
Best Local Similarity	83.3%;	Pred. No. 46;		
Matches	5;	Conservative	1;	Mismatches 0;
				Indels 0;
				Gaps 0;

QY 3 wkgyw 8
||:|||
Db 329 WGRGYW 334

RESULT 7
 AC2394
 hypothetical protein asr4707 [imported] - Anabaena sp. (strain PCC 7120)
 C:Species: Anabaena sp.
 A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
 C:Accession: AC2394
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriqichh
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A:Reference numbers: AB1607; MUID:21595285; PMID:11759640
 A:Accession: AC2394
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-80 <KUR>
 A:Cross-references: GB:BA000013; PIDN:BA876406.1; PID:g17133844; GSPDB:GN00179
 A:Experimental source: Strain PCC 7120
 C:Genetics:
 C:Gene: asr4707

Query Match 71.2%; Score 42; DB 2; Length 80;
Best Local Similarity 85.7%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY 1 gYwgkgy 7
Db 38 GYWGKSY 44

RESULT 8
D83812
hypothetical protein BH1300 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: D83812
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiro
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: AB3650; MUID:20512582; PMID:11058132
A:Accession: D83812

Query Match

71.28; Score 42; DB 2; Length 196;

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-180 <STO>
A;Cross-references: GB;
A;Experimental source:
C;Genetics:
A;Gene: BH1300

Query Match 71.2%; Score 42; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWGKY 7
Db 102 YWGKY 107

RESULT 9

Hypothetical protein BH2121 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2000
C:Accession: A83915
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; F.
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: A83915
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-181 <STO>
A:Cross-references: GB:AP001514; GB:BA000004; NID:g10174613; PIDN:BA005840
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH2121

Query Match 71.2%; Score 42; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ywgky 7
Db 100 YWGGY 105

RESULT	10
--------	----

acetyltransferase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AE2918
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen,
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li,
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Go-
ster, E.W.

A; Reference number: AB2577; PMID:11743193

A;Accession: M22210
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-196 <KUR>
A;Cross-references: GB
A;Experimental source:
C;Genetics:
A;Gene: Atu2782
A;Map position: circular

71.28; Score 42; DB 2; Length 196;

```

Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ywgkgy 7
|||||
Db 116 YWGRGY 121

RESULT 11
H97692
conserved hypothetical protein BH0837 (AP001510) [imported] - Agrobacterium tumefaciens
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: H97692
R:Goodner, B.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
A.: Liu, F.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; PMID:11743194
A:Accession: H97692
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-207 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK88497.1; PID:g15158008; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_5049
A:Map position: circular chromosome

Query Match 71.2%; Score 42; DB 2; Length 207;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ywgkgy 7
|||||
Db 127 YWGRGY 132

RESULT 12
A69491
probable formylmethanofuran dehydrogenase (EC 1.2.99.5) (molybdenum/tungsten) chain A -
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C:Accession: A69491
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glocke, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; PMID:98049343
A:Accession: A69491
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-563 <KLE>
A:Cross-references: GB:AE000970; NID:g2689293; PIDN:AA889325.1; PID:g264861
C:Superfamily: formylmethanofuran dehydrogenase (molybdenum) alpha chain
C:Keywords: iron-sulfur protein; molybdenum; molybdopterin; oxidoreductase; tungsten

Query Match 71.2%; Score 42; DB 2; Length 563;
Best Local Similarity 71.4%; Pred. No. 75;
Matches 5; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 2 ywgkgyw 8
|||||
Db 511 YWGRTYW 517

RESULT 13
A56390
mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - Strept
C:Species: Streptococcus pneumoniae
C:Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 15-Oct-1999
C:Accession: A56390
R:Clarke, V.A.; Platt, N.; Butters, T.D.
J. Biol. Chem. 270, 8805-8814, 1995
A:Title: Cloning and expression of the beta-N-acetylglucosaminidase gene from Strepto
A:Reference number: A56390; MUID:95238375
A:Accession: A56390
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1311 <CLA>
A:Cross-references: GB:L36923; NID:g784896; PIDN:AA41450.1; PID:g784897
C:Genetics:
A:Gene: strH
C:Keywords: glycosidase; hydrolase; tandem repeat
F:1-33/Domain: signal sequence #status predicted <SIG>

Query Match 71.2%; Score 42; DB 2; Length 1312;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywgkgyw 8
|||||
Db 870 YWSKGW 876

RESULT 14
E95006
beta-N-acetylhexosaminidase [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: E95006
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzappl
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A:Title: Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: E95006
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1312 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK74246.1; PID:g14971522; GSPDB:GN00164; TIGR:
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0057

Query Match 71.2%; Score 42; DB 2; Length 1312;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywgkgyw 8
|||||
Db 871 YWSKGW 877

RESULT 15
A97879
beta-N-acetylhexosaminidase (EC 3.2.1.52) [imported] - Streptococcus pneumoniae (stra
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C:Accession: A97879
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A97879

```


A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1312 <KUR>
A;Cross-references: GB:AE007317; PIDN:AAK98861.1; PID:gl5457590; GSPDB:GN00174
C;Genetics:
A;Gene: strH
C;Keywords: glycosidase; hydrolase

Query Match 71.2%; Score 42; DB 2; Length 1312;
Best Local Similarity 71.4%; Pred. NO. 1.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywgkgyw 8
|| ||:
Db 871 YWSKGWW 877

Search completed: August 12, 2002, 17:22:44
Job time: 181 sec


```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF312031; AAG27731.1; -.
DR InterPro; IPR002528; UPF0013.
DR Pfam; PF01554; UPF0013. 2.
KW Transport; Sodium transport; Inner membrane.
FT TRANSMEM 56 76
FT TRANSMEM 99 119
FT TRANSMEM 147 167
FT TRANSMEM 172 192
FT TRANSMEM 202 222
FT TRANSMEM 247 267
FT TRANSMEM 280 300
FT TRANSMEM 325 345
FT TRANSMEM 361 381
FT TRANSMEM 402 422
FT TRANSMEM 430 450
SQ SEQUENCE 462 AA; 47860 MW; 0CAF82EC20652D11 CRC64;

Query Match          69.5%; Score 41; DB 1; Length 462;
Best Local Similarity 58.3%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

QY 1 gywg----kgwv 8
DB 408 GYWGIGFPTGYW 419

RESULT 4
ID DNL4_HUMAN STANDARD; PRT; 844 AA.
AC P49517.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNL4 ligase IV (EC 6.5.1.1) (Polydeoxyribonucleotide synthase [ATP]).
GN LIG4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Prostate;
RX MEDLINE=95280920; PubMed=7760816;
RA Wei Y.-F., Robins P., Carter K., Caldecott K., Pappin D.J.C.,
RA Yu G.-L., Wang R.-P., Shell B.K., Nash R.A., Schar P., Barnes D.E.,
RA Haseltine W.A., Lindahl T.;
RT "Molecular cloning and expression of human cDNAs encoding a novel DNA
RT ligase IV and DNA ligase III, an enzyme active in DNA repair and
RT recombination.";
RL Mol. Cell. Biol. 15:3206-3216(1995).
CC -!- CATALYTIC ACTIVITY: ATP + {deoxyribonucleotide}(N) +
CC {deoxyribonucleotide}(M) = AMP + diphosphate +
CC {deoxyribonucleotide}(N+M).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: TESTIS, THYMUS, PROSTATE AND HEART.
CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT DNA LIGASE FAMILY.
CC -!- SIMILARITY: CONTAINS 2 BCT DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X83441; CAA58467.1; -.
DR MIM; 601837; -.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR000977; DNA_ligase.
```

```
DR Pfam; PF00533; BRCT; 2.
DR Pfam; PF01088; DNA_ligase; 1.
DR SMART; SM00292; BRCT; 2.
DR PROSITE; PS01072; BRCT; 2.
DR PROSITE; PS00697; DNA_LIGASE_A1; 1.
DR PROSITE; PS00333; DNA_LIGASE_A2; 1.
DR PROSITE; PS01060; DNA_LIGASE_A3; 1.
KW DNA repair; DNA replication; DNA recombination; Cell division; Ligase;
KW ATP-binding; Nuclear protein; Repeat.
FT DOMAIN 587 676
FT DOMAIN 741 844
FT BINDING 206 206 AMP (BY SIMILARITY).
SQ SEQUENCE 844 AA; 96154 MW; D4BE16F211A3BE97 CRC64;

Query Match          69.5%; Score 41; DB 1; Length 844;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gywgk 6
DB 402 GYWGKG 407

RESULT 5
ID UGGG_SCHPO STANDARD; PRT; 1448 AA.
AC Q09140;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE UDP-glucose-glycoprotein glucosyltransferase precursor (EC 2.4.1.-)
DE (UDP-Glc-glycoprotein glucosyltransferase).
GN GPT1 OR SFBPJ4664.06.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
[1]
RN SEQUENCE FROM N.A., SEQUENCE OF 19-31; 148-158; 401-412; 557-568;
RP 723-730; 983-990 AND 1118-1130, FUNCTION, SUBCELLULAR LOCATION,
RP COFACTOR, SUBUNIT, AND GLYCOSYLATION.
RX MEDLINE=96181349; PubMed=8631292;
RA Fernandez F., Jannatipour M., Hellman U., Rokeach L.A., Parodi A.J.;
RT "A new stress protein: synthesis of Schizosaccharomyces pombe
RT UDP-Glc-glycoprotein glucosyltransferase mRNA is induced by stress
RT conditions but the enzyme is not essential for cell viability.";
RL EMBL J. 15:705-713(1996).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=972;
RA Wood V., Warren T., Harris D., Barrell B.G., Rajandream M.A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE OF 19-31, AND CHARACTERIZATION.
RX MEDLINE=95074086; PubMed=7982990;
RA Fernandez F.S., Trombetta S.E., Hellman U., Parodi A.J.;
RT "Purification to homogeneity of UDP-glucose:glycoprotein
RT glucosyltransferase from Schizosaccharomyces pombe and apparent
RT absence of the enzyme from Saccharomyces cerevisiae.";
RL J. Biol. Chem. 269:30701-30706(1994).
CC -!- FUNCTION: UNFOLDED, DENATURED GLYCOPROTEINS ARE SUBSTANTIALLY
CC BETTER SUBSTRATES FOR GLUCOSYLATION BY THIS ENZYME THAN ARE THE
CC CORRESPONDING NATIVE PROTEINS. THIS PROTEIN AND TRANSIENT
CC GLUCOSYLATION MAY BE INVOLVED IN MONITORING AND/OR ASSISTING THE
CC FOLDING AND ASSEMBLY OF NEWLY MADE GLYCOPROTEINS, IN ORDER TO
CC IDENTIFY GLYCOPROTEINS THAT NEED ASSISTANCE IN FOLDING FROM
CC CHAPERONES.
CC -!- COFACTOR: REQUIRES CALCIUM AND MANGANESE IONS FOR ACTIVITY.
CC -!- PATHWAY: GLYCOSYLATION.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
```


Query Match 67.8%; Score 40; DB 1; Length 518;
Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 wgkgyw 8
DB 491 WGPCYW 496

RESULT 7
SAP_RAT STANDARD; PRT; 554 AA.
AC P10960; Q62841; Q64190;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (prosaposin).
GN PSAP OR SGPI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Sertoli cells;
RX MEDLINE=89000647; PubMed=3048385;
RA Collard M.W., Sylvester S.R., Tsuruta J.K., Griswold M.D.;
RT "Biosynthesis and molecular cloning of sulfated glycoprotein 1
secreted by rat Sertoli cells: sequence similarity with the
70-kilodalton precursor to sulfatide/GM1 activator.";
RL Biochemistry 27:4557-4564(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96128341; PubMed=8573994;
RA Morales C.R., El-Alfy M., Zhao Q., Igdoura S.A.;
RT "Molecular role of sulfated glycoprotein-1 (SGP-1/prosaposin) in
Sertoli cells.";
RL Histol. Histopathol. 10:1023-1034(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RX MEDLINE=96175245; PubMed=8601692;
RA Morales C.R., El-Alfy M., Zhao Q., Igdoura S.A.;
RT "Expression and tissue distribution of rat sulfated glycoprotein-1
(prosaposin).";
RL J. Histochem. Cytochem. 44:327-337(1996).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: CONTAINS 2 SAPOSIN A-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 4 SAPOSIN B-TYPE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M19936; AAA42136.1; -;
CC EMBL; S81353; AAB36042.2; -;
CC EMBL; S81373; AAB36233.2; -;
CC PIR; A28716; A28716.
CC InterPro; IPR003119; Sapa.
CC InterPro; IPR000004; SApB.
CC InterPro; IPR003259; Saposin.
CC Pfam; PF02199; SAPA; 2.
CC ProDom; PD001732; SApB; 3.
CC ProDom; PD012321; Saposin; 1.
CC SMART; SM00162; SAPA; 2.
CC SMART; SM00118; SApB; 4.

KW Sulfation; Signal; Glycoprotein; Repeat.
FT SIGNAL 1 16
FT CHAIN 17 554 SULFATED GLYCOPROTEIN 1.
FT DOMAIN 21 54 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 138 138 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 193 274 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 310 391 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 435 516 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 521 554 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 63 138 BY SIMILARITY.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 197 270 BY SIMILARITY.
FT DISULFID 200 264 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 314 387 BY SIMILARITY.
FT DISULFID 317 381 BY SIMILARITY.
FT DISULFID 345 356 BY SIMILARITY.
FT DISULFID 439 512 BY SIMILARITY.
FT DISULFID 442 506 BY SIMILARITY.
FT DISULFID 470 481 BY SIMILARITY.
FT CARBOHYD 80 80 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 115 115 P -> L (IN REF. 2 AND 3).
FT CONFLICT 299 299 D -> E (IN REF. 2).
FT CONFLICT 462 462 I -> V (IN REF. 3).
FT CONFLICT 527 527 W -> R (IN REF. 3).
FT CONFLICT 536 536 S -> M (IN REF. 3).
SQ SEQUENCE 554 AA; 61123 MW; DFE3F3A0520C6B CRC64;

Query Match 67.8%; Score 40; DB 1; Length 554;
Best Local Similarity 83.3%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 wgkgyw 8
DB 527 WGPCYW 532

RESULT 8
ACM3_MOUSE STANDARD; PRT; 589 AA.
ID ACM3_MOUSE
AC Q9ERZ3; Q64055;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Muscarinic acetylcholine receptor M3 (Mm3 mAChR).
GN CHRM3 OR CHRM-3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Gomez J., Weiss J.;
RT "Isolation, sequence and functional expression of mouse muscarinic
acetylcholine receptor genes.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 314-439 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95179320; PubMed=7874308;
RA Andre C., Dos Santos G., Koulakoff A.;
RT "Cultured neurons from mouse brain reproduce the muscarinic receptor
profile of their tissue of origin.";
RL Eur. J. Neurosci. 6:1691-1701(1994).
CC -1- FUNCTION: THE MUSCARINIC ACETYLCHOLINE RECEPTOR MEDIATES VARIOUS
CELLULAR RESPONSES, INCLUDING INHIBITION OF ADENYLATE CYCLASE
BREAKDOWN OF PHOSPHOLIPIDS & MODULATION OF POTASSIUM CHANNELS
THROUGH THE ACTION OF G PROTEINS. PRIMARY TRANSDUCING EFFECT IS PI

KW Postsynaptic membrane; Ionic channel; Glycoprotein; Transmembrane;
KW Phosphorylation; Multigene family; G-protein coupled receptor.
FT DOMAIN 1 66
FT TRANSSEM 67 90
FT DOMAIN 91 103
FT TRANSSEM 104 124
FT DOMAIN 125 141
FT TRANSSEM 142 163
FT DOMAIN 164 183
FT TRANSSEM 184 206
FT DOMAIN 207 228
FT TRANSSEM 229 251
FT DOMAIN 252 491
FT TRANSSEM 492 512
FT DOMAIN 513 526
FT TRANSSEM 527 546
FT DOMAIN 547 589
FT CARBOHYD 6 6
FT CARBOHYD 15 15
FT CARBOHYD 41 41
FT CARBOHYD 48 48
FT CARBOHYD 52 52
FT DISULFID 140 220
FT CONFLICT 184 184
FT CONFLICT 516 516
FT CONFLICT 556 556
FT SEQUENCE 589 AA; 66065 MW; 9A5EF2FA653C830A CRC64;

Query Match 67.8%; Score 40; DB 1; Length 589;
Best Local Similarity 71.4%; Pred. No. 38;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ywkgqyw 8
DB 524 YWNLGYW 530

RESULT 10
ACM3_BOVIN STANDARD; PRT; 590 AA.
ID ACM3_BOVIN
AC P41384;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Muscarinic acetylcholine receptor M3.
GN CHRM3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=94339178; PubMed=8061048;
RA Lee P.H., Hodges P.K., Glickman E., Chang K.J.;
RT "Cloning and expression of a cDNA encoding bovine muscarinic
acetylcholine m3 receptor.";
RL Biochim. Biophys. Acta 1223:151-154(1994).
RN [2]
RP SEQUENCE OF 327-467 FROM N.A.
RC TISSUE=Adrenal gland;
RA Sui A.-L., Chou W.-Y., Kao L.-S.;
RL Submitted (XXA-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE MUSCARINIC ACETYLCHOLINE RECEPTOR MEDIATES VARIOUS
CELLULAR RESPONSES, INCLUDING INHIBITION OF ADENYLATE CYCLASE,
BREAKDOWN OF PHOSPHOINOSITIDES & MODULATION OF POTASSIUM CHANNELS
THROUGH THE ACTION OF G PROTEINS. PRIMARY TRANSDUCING EFFECT IS P1
TURNOVER.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U08286; AAA51866.1; -;
DR EMBL; L27103; AAA30653.1; -;
DR GCRDb; GCR_0849; -;
DR GCRDb; GCR_1227; -;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1;
DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2;
DR PROSITE; PS50262; G_PROTEIN_RECP_F1_3;
KW Postsynaptic membrane; Ionic channel; Glycoprotein; Transmembrane;
KW Phosphorylation; Multigene family; G-protein coupled receptor.
FT DOMAIN 1 67
FT TRANSSEM 68 91
FT DOMAIN 92 104
FT TRANSSEM 105 125
FT DOMAIN 126 142
FT TRANSSEM 143 164
FT DOMAIN 165 184
FT TRANSSEM 185 207
FT DOMAIN 208 229
FT TRANSSEM 230 252
FT DOMAIN 253 492
FT TRANSSEM 493 513
FT DOMAIN 514 527
FT TRANSSEM 528 547
FT DOMAIN 548 590
FT CARBOHYD 6 6
FT CARBOHYD 7 7
FT CARBOHYD 15 15
FT CARBOHYD 41 41
FT CARBOHYD 48 48
FT CARBOHYD 53 53
FT DISULFID 141 221
FT CONFLICT 424 424
FT CONFLICT 438 438
FT CONFLICT 440 440
FT CONFLICT 452 452
FT CONFLICT 461 461
FT CONFLICT 467 467
FT SEQUENCE 590 AA; 66103 MW; 4DE04EDE33CCA8D6 CRC64;
SQ
Query Match 67.8%; Score 40; DB 1; Length 590;
Best Local Similarity 71.4%; Pred. No. 38;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 ywkgqyw 8
DB 525 YWNLGYW 531
RESULT 11
ACM3_PIG STANDARD; PRT; 590 AA.
ID ACM3_PIG
AC P11483;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Muscarinic acetylcholine receptor M3.
GN CHRM3.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]

```
RP SEQUENCE FROM N.A.
RX MEDLINE=88296835; PubMed=3402600;
RA Akiba I., Kubo T., Maeda A., Bujo H., Nakai J., Mishina M., Numa S.;
RT "Primary structure of porcine muscarinic acetylcholine receptor III
RL and antagonist binding studies.";
RT FEBS Lett. 235:257-261(1988).
CC -!- FUNCTION: THE MUSCARINIC ACETYLCHOLINE RECEPTOR MEDIATES VARIOUS
CC CELLULAR RESPONSES, INCLUDING INHIBITION OF ADENYLATE CYCLASE,
CC BREAKDOWN OF PHOSPHOINOSITIDES & MODULATION OF POTASSIUM CHANNELS
CC THROUGH THE ACTION OF G PROTEINS. PRIMARY TRANSDUCING EFFECT IS PI
CC TURNOVER.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X12712; CAA31215.1; -.
CC PIR; S01114; S01114.
CC GCRDb; GCR_0104; -.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm.1; 1.
CC PRINTS; PR00237; GPCRHHODPSN.
CC PROSITE; PS00237; G-PROTEIN_RECP_FL_1; 1.
CC PROSITE; PS0262; G-PROTEIN_RECP_FL_2; 1.
CC Postsynaptic membrane; Ionic channel; Glycoprotein; Transmembrane;
CC Phosphorylation; Multigene family; G-protein coupled receptor.
KW DOMAIN 1 67
FT TRANSMEM 68 91
FT DOMAIN 92 104
FT DOMAIN 105 125
FT TRANSMEM 126 142
FT TRANSMEM 143 164
FT DOMAIN 165 184
FT TRANSMEM 185 207
FT TRANSMEM 208 229
FT TRANSMEM 230 252
FT TRANSMEM 253 492
FT TRANSMEM 493 513
FT TRANSMEM 514 527
FT TRANSMEM 528 547
FT TRANSMEM 548 590
FT CARBOHYD 6 6
FT CARBOHYD 7 7
FT CARBOHYD 8 8
FT CARBOHYD 15 15
FT CARBOHYD 41 41
FT CARBOHYD 48 48
FT CARBOHYD 53 53
FT DISULFID 141 221
FT SEQUENCE 590 AA; 66077 MW; 99980D2A4802FD32A CRC64;
SQ
Query Match 67.8%; Score 40; DB 1; Length 590;
Best Local Similarity 71.4%; Pred. No. 38;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 ywgkgyw 8
DB 525 YWNLGYW 531
RESULT 12
ID YD35_MYCPN STANDARD; PRT; 741 AA.
AC P75443;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT
```

```
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein MPN335 (F10_orf741).
GN MPN335 OR MP501.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Piagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO M.PNEUMONIAE MPN333.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF000049; AAB96149.1; -.
CC InterPro; IPR000188; GABAA_receptor.
CC Pfam; PF02932; Neur_chan_memb; 1.
CC KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 34 54
FT TRANSMEM 54 96
FT TRANSMEM 120 140
FT TRANSMEM 156 176
FT TRANSMEM 187 207
FT TRANSMEM 655 675
FT TRANSMEM 715 735
FT SEQUENCE 741 AA; 85025 MW; 1776A96BCF83567F CRC64;
SQ
Query Match 67.8%; Score 40; DB 1; Length 741;
Best Local Similarity 71.4%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 ywgkgyw 8
DB 280 YWNLGYW 286
RESULT 13
ID YM05_MARPO STANDARD; PRT; 228 AA.
AC P38453;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 26.5 kDa protein in RPS2 3'region (ORF 228).
GN YMF5.
OS Marchantia polymorpha (Liverwort).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;
OC Marchantiales; Marchantiaceae; Marchantia.
OX NCBI_TaxID=3197;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92114051; PubMed=1731062;
RA Oda K., Yamato K., Ohta E., Nakamura Y., Takemura M., Nozato N.,
RA Akashi K., Kanegae T., Ogura Y., Kohchi T., Ohyama K.;
RT "Gene organization deduced from the complete sequence of liverwort
RT Marchantia polymorpha mitochondrial DNA. A primitive form of plant
RT mitochondrial genome.";
RL J. Mol. Biol. 223:1-7(1992).
RL
```

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M68929; AAC09463.1; -.
DR PIR: S26009; S26009.
DR Mendel; 2116; MARPO:ymf5;1.
DR InterPro; IPR002541; CytC_asm.
DR Pfam; PF01578; CytC_asm; 1.
DR PRINTS; PR01386; CMCBIOGNSIS.
KW Mitochondrion; Hypothetical protein.
SQ SEQUENCE 228 AA; 26461 MW; BF13C3616D6C7D35 CRC64;

Query Match 66.1%; Score 39; DB 1; Length 228;
Best Local Similarity 62.5%; Pred. No. 22;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 gwgkgyw 8
DB 109 GFWGRPMW 116
I:|||||
|:|||||

RESULT 14
ID INAL_TRIHA STANDARD; PRT; 573 AA.
AC P34054;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Amino-acid permease INDAL.
GN INDAL.
OS Trichoderma harzianum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Mitosporic Hypocreales; Trichoderma.
OX NCBI_TaxID=5544;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IMI 206040;
RX MEDLINE=95291429; PubMed=7773384;
RA Vasseur V.V., van Montagu M.M., Goldman G.G.H.;
RT "Trichoderma harzianum genes induced during growth on Rhizoctonia
RT solani cell walls.";
RL Microbiology 141:767-774(1995).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- INDUCTION: DURING MYCOPARASITISM.
CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z22594; CAA80308.1; -.
DR PIR: S33212; S33212.
DR InterPro; IPR002293; AA_rel_permease_1.
DR InterPro; IPR002027; Amino_acid_permease.
DR Pfam; PF00324; aa_permeases; 1.
DR PROSITE; PS00218; AMINO-ACID_PERMEASE_1; 1.
KW Transport; Amino-acid transport; Transmembrane.
FT TRANSMEM 72 92 POTENTIAL.
FT TRANSMEM 99 117 POTENTIAL.
FT TRANSMEM 176 200 POTENTIAL.
FT TRANSMEM 212 229 POTENTIAL.
```

```
FT TRANSMEM 257 280 POTENTIAL.
FT TRANSMEM 296 315 POTENTIAL.
FT TRANSMEM 351 371 POTENTIAL.
FT TRANSMEM 398 415 POTENTIAL.
FT TRANSMEM 425 444 POTENTIAL.
FT TRANSMEM 470 494 POTENTIAL.
FT TRANSMEM 511 527 POTENTIAL.
SQ SEQUENCE 573 AA; 62850 MW; 5FB0A806934DB55D CRC64;

Query Match 66.1%; Score 39; DB 1; Length 573;
Best Local Similarity 71.4%; Pred. No. 52;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ywgkgyw 8
DB 239 YWGARYW 245
I:|||||
|:|||||

RESULT 15
ID CDGT_KLEPN STANDARD; PRT; 655 AA.
AC P08704;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Cyclomaltodextrin glucanotransferase precursor (EC 2.4.1.19)
DE (Cyclomaltodextrin-glycosyltransferase) (CGTase).
GN CGT.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M5A1;
RX MEDLINE=87163498; PubMed=2951300;
RA Binder F., Huber O., Boeck A.;
RT "Cyclodextrin-glycosyltransferase from Klebsiella pneumoniae M5a1:
RT cloning, nucleotide sequence and expression.";
RL Gene 47:269-277(1986).
CC -1- CATALYTIC ACTIVITY: Degrades starch to cyclodextrins by formation
CC of a 1,4-alpha-D-glucosidic bond.
CC -1- COFACTOR: BINDS TWO CALCIUM IONS.
CC -1- SUBUNIT: MONOMER.
CC -1- MISCELLANEOUS: CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE
CC IN THE AMINO-TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND
CC IN STARCH, AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER
CC ACTIVITIES, INCLUDING THE RECONSTITUTION OF AN
CC ALPHA-1,4-GLUCOSIDIC LINKAGE FOR CYCLIZING THE
CC MALTOOLIGOSACCHARIDE PRODUCED.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M15264; AAA25059.1; -.
DR PIR: A29023; ALKRG.
DR HSSP; P30920; 1CGT.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR002044; CBD_4.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02806; alpha-amylase_C; 1.
DR Pfam; PF00686; CBD_4; 1.
DR ProDom; PD001568; CBD_4; 1.
KW Transferase; Glycosyltransferase; Calcium; Signal.
FT SIGNAL 1 30
```

FT CHAIN 31 655 CYCLOMALTODEXTRIN GLUCANOTRANSFERASE.
FT ACT_SITE 253 253 BY SIMILARITY.
FT ACT_SITE 287 287 BY SIMILARITY.
FT ACT_SITE 363 363 BY SIMILARITY.
SQ SEQUENCE 655 AA; 73024 MW; DB8F26332BED26A7 CRC64;

Query Match 66.1%; Score 39; DB 1; Length 655;
Best Local Similarity 62.5%; Pred. No. 59;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gywgkgyw 8
|||:|:
Db 122 GYWGGRDYF 129

Search completed: August 12, 2002, 17:27:25
Job time: 337 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:22:13 ; Search time 41.64 Seconds

(without alignments)
33.236 Million cell updates/sec

Title: 10-071247-1

Perfect score: 59

Sequence: 1 gywgkgyw 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	79.7	748	17 Q9HLG7	Q9hlg7 thermoplasma
2	45	76.3	441	16 Q9KFW3	Q9kfw3 bacillus ha
3	45	76.3	574	4 Q9H5J0	Q9h5j0 homo sapien
4	44	74.6	198	10 Q9AXA8	Q9axa8 oryza sativ
5	44	74.6	716	17 Q9HLF3	Q9hlf3 thermoplasma
6	43.5	73.7	83	16 Q34654	Q34654 bacillus su
7	43.5	73.7	187	16 Q9HV15	Q9hvl5 pseudomonas
8	43	72.9	277	2 Q9F3T0	Q9f3t0 rhodothermu
9	43	72.9	358	16 Q9I5J3	Q9i5j3 pseudomonas
10	43	72.9	470	17 Q9VAX5	Q9vax5 aeropyrum p
11	42	71.2	126	5 Q9VNF1	Q9vnf1 drosophila
12	42	71.2	133	9 Q21906	Q21906 bacterioph
13	42	71.2	180	16 Q9KDB3	Q9kdb3 bacillus ha
14	42	71.2	181	16 Q9KBL5	Q9kbl5 bacillus ha
15	42	71.2	446	2 Q9F189	Q9f189 alcaligenes
16	42	71.2	563	17 Q28349	Q28349 archaeoglob

17	41	69.5	105	12	Q9BWT5	Q9emt5 ansacta moo
18	41	69.5	145	16	Q9RT39	Q9rt39 deinococcus
19	41	69.5	206	2	O05148	O05148 rhodococcus
20	41	69.5	912	13	Q9QYB1	Q9qyb1 gallus gall
21	40	67.8	252	2	P95462	P95462 plectonema
22	40	67.8	469	10	Q9FTG0	Q9ftg0 oryza sativ
23	40	67.8	489	16	P73420	P73420 synechocyst
24	39	66.1	102	17	Q9FEW6	Q9feyw6 aeropyrum p
25	39	66.1	137	3	Q01150	Q01150 magnaporthe
26	39	66.1	138	3	Q01152	Q01152 magnaporthe
27	39	66.1	159	16	Q9CN19	Q9cn19 pasteurella
28	39	66.1	177	16	Q9ABBA	Q9abb4 caulobacter
29	39	66.1	186	16	Q9TSA7	Q9tsa7 streptococc
30	39	66.1	247	10	Q9LZZ1	Q9lzz1 arabidopsis
31	39	66.1	270	16	Q9HZ21	Q9hzz21 pseudomonas
32	39	66.1	360	2	O33405	O33405 pseudomonas
33	39	66.1	396	10	O49647	O49647 arabidopsis
34	39	66.1	499	2	Q9S2D9	Q9s2d9 streptomyce
35	39	66.1	736	2	Q9L053	Q9l053 streptomyce
36	39	66.1	961	3	Q12361	Q12361 saccharomyc
37	38.5	65.3	89	2	Q9ZNI3	Q9zni3 pseudomonas
38	38.5	65.3	273	16	Q930Z1	Q930z1 rhizobium m
39	38	64.4	105	11	O08939	O08939 meriones un
40	38	64.4	115	10	Q9LN41	Q9ln41 arabidopsis
41	38	64.4	117	10	Q93WK6	Q93wk6 arabidopsis
42	38	64.4	132	17	Q9HM86	Q9hm86 halobacteri
43	38	64.4	138	11	O35224	O35224 cynomys lud
44	38	64.4	163	17	Q96ZX8	Q96zxx8 sulfolobus
45	38	64.4	172	11	Q9WUJ4	Q9wuuj4 mus musculu

ALIGNMENTS

RESULT 1

Q9HLG7 ID Q9HLG7 PRELIMINARY; PRT; 748 AA.
AC Q9HLG7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL PROTEIN TA0261.
TA0261.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmataceae;
OC Thermoplasma
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumelster W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum.";
RL Nature 407:508-513(2000).
DR EMBL; AL445063; CAC11406.1; .
KW Hypothetical protein: Complete proteome.
SQ SEQUENCE 748 AA; 82250 MW; 73A030A0E217EE39 CRC64;

Query Match 79.7%; Score 47; DB 17; Length 748;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywgkgyw 8
||| |
Db 532 YWGKAYW 538

RESULT 2

Q9KFW3 ID Q9KFW3 PRELIMINARY; PRT; 441 AA.

Qy 1 gywkgyw 8
|:|||||

QY 2 YWgkgyw 8
|||: ||
Db 498 YWGRSYW 504

```

RESULT 6
O34654 ID O34654 PRELIMINARY; PRT; 83 AA.
AC O34654;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE YOLA.
GN YOLA OR YODI.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Ghim S.-Y., Jeong Y.-M., Choi S.-K., Park S.-H.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Wambutt R., Wedler H., Lapidus A., Sorokin A., Ehrlich D.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Bouilliet S., Brusier C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.-K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.-Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presceac E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivoita C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF006665; AAB81166.1; -.
DR EMBL; AF015775; AAE72056.1; -.
DR EMBL; Z99114; CAB13852.1; -.
KW Complete proteome.
SQ SEQUENCE 83 AA; 9194 MW; 99F58EA2F0F36A43 CRC64;

Query Match 73.7%; Score 43.5; DB 16; Length 83;
Best Local Similarity 87.5%; Pred. No. 8.4;

Matches 7; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 gywgkgyw 8
DB 53 GYWG-GYW 59
|||||

RESULT 7
Q9HV15 ID Q9HV15 PRELIMINARY; PRT; 187 AA.
AC Q9HV15;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE HYPOTHETICAL PROTEIN PA4793.
GN PA4793.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
DR EMBL; AB004892; AAG08179.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 187 AA; 21281 MW; 8908E9EBEC51897 CRC64;

Query Match 73.7%; Score 43.5; DB 16; Length 187;
Best Local Similarity 87.5%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 gywgkgyw 8
DB 109 GYWG-GYW 115
|||||

RESULT 8
Q9F3T0 ID Q9F3T0 PRELIMINARY; PRT; 277 AA.
AC Q9F3T0;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHETICAL 31.4 KDA PROTEIN.
OS Rhodothermus marinus (Rhodothermus obamensis).
OC Bacteria; CFB group; Rhodothermus group; Rhodothermus.
OX NCBI_TaxID=29549;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PRQ- 62B;
RX MEDLINE=20576172; PubMed=11133964;
RA Santana M., Pereira M.M., Elias N.P., Soares C.M., Teixeira M.;
RT "Gene cluster of rhodothermus marinus high-potential iron-sulfur
RT protein: oxygen oxidoreductase, a cca3-type oxidase belonging to the
RT superfamily of heme-copper oxidases."
RL J. Bacteriol. 183:687-699(2001).
DR EMBL; AJ249578; CAC08530.1; -.
KW Hypothetical protein.
SQ SEQUENCE 277 AA; 31355 MW; 73CF0EF4970E1E7A CRC64;

Query Match 72.9%; Score 43; DB 2; Length 277;

```

Best Local Similarity 71.4%; Pred. No. 35; Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;		RL DNA Res. 6:83-101(1999). DR EMBL; AP000062; BAA80823.1; -. DR InterPro; IPR000712; Bcl_2. DR InterPro; IPR000447; FAD_gly3P_dh. DR InterPro; IPR000205; NAD_binding. DR PRINTS; PR01001; FADG3PDH. DR PROSITE; PS01258; BH2; 1. KW Hypothetical protein; Complete proteome. SQ SEQUENCE 470 AA; 51138 MW; 27CA636B23A7FF7B CRC64;	
Query Match 72.9%; Score 43; DB 17; Length 470; Best Local Similarity 83.3%; Pred. No. 62; Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;		QY 3 wdgkgyw 8 DB 329 WGRGYW 334	
RESULT 11		Q9VNF1 PRELIMINARY; PRT; 126 AA.	
ID	Q9VNF1	AC Q9VNF1	
DT	01-MAY-2000 (TRENBLrel. 13, Created)	DT 01-MAY-2000 (TRENBLrel. 13, Created)	
DT	01-MAY-2000 (TRENBLrel. 13, Last sequence update)	DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)	
DE	CG14673 PROTEIN.	DE CG14673 PROTEIN.	
GN	CG14673.	GN CG14673.	
OS	Drosophila melanogaster (Fruit fly).	OS Drosophila melanogaster (Fruit fly).	
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC	Ephydroidea; Drosophilidae; Drosophila.	OC Ephydroidea; Drosophilidae; Drosophila.	
OX	NCBI_TaxID=7227;	OX NCBI_TaxID=7227;	
RN	[1]	RN [1]	
RP	SEQUENCE FROM N.A.	RP SEQUENCE FROM N.A.	
RC	STRAIN=BERKELEY.	RC STRAIN=BERKELEY.	
RX	MEDLINE=20196006; PubMed=10731132;	RX MEDLINE=20196006; PubMed=10731132;	
RA	Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,	RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,	
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,	RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,	
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,	RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,	
RA	Sutton G.G., Wortman J.R., Yeandle M.D., Zhang Q., Chen L.X.,	RA Sutton G.G., Wortman J.R., Yeandle M.D., Zhang Q., Chen L.X.,	
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,	RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,	
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,	RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,	
RA	April J.F., Agayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,	RA April J.F., Agayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,	
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,	RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,	
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,	RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,	
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,	RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,	
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,	RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,	
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,	RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,	
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,	RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,	
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,	RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,	
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,	RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,	
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,	RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,	
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,	RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,	
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,	RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,	
RA	Hoskin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,	RA Hoskin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,	
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,	RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,	
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,	RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,	
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,	RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,	
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,	RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,	
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,	RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,	
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,	RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,	
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,	RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,	
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,	RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,	
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,	RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,	
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,	RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,	
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,	RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,	
RA	Svirskas R., Turner C., Turner R., Venter E., Wang A.H., Wang X.,	RA Svirskas R., Turner C., Turner R., Venter E., Wang A.H., Wang X.,	
RA	Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,	RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,	
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,	RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,	
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,	RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,	
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,	RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,	

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL: AF003602; AAF51986.1; -;
 DR Flybase; FBgn0037352; C014673.
 SQ SEQUENCE 126 AA; 14140 MW; DC68AB3DF1F5F820 CRC64;

Query Match 71.2%; Score 42; DB 5; Length 126;
 Best Local Similarity 71.4%; Pred. No. 22;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ywgkgyw 8
 || |||
 Db 37 YWNSGYW 43

RESULT 12

O21906 PRELIMINARY; PRT; 133 AA.
 AC O21906;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 15.7 KDA PROTEIN.
 OS Bacteriophage sk1.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
 OX NCBI_TaxID=31532;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98043537; PubMed=9383189;
 RA Chandry P.S., Moore S.C., Boyce J.D., Davidson B.E., Hillier A.J.;
 RT "Analysis of the DNA sequence, gene expression, origin of replication
 and modular structure of the *Lactococcus lactis* lytic bacteriophage
 sk1."
 RL Mol. Microbiol. 26:49-64(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Chandry P.S., Moore M.C., Boyce J.D., Davidson B.E., Hillier A.J.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF011378; AAB70078.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 133 AA; 15682 MW; 7D985B5C51E961BC CRC64;

Query Match 71.2%; Score 42; DB 9; Length 133;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ywgkgy 7
 |||||
 Db 102 YWGKGY 107

RESULT 13

O9KDB3 PRELIMINARY; PRT; 180 AA.
 AC O9KDB3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE BH1300 PROTEIN.
 GN BH1300.
 OS *Bacillus halodurans*.
 OC Bacteria; Firmicutes; *Bacillus*/Clostridium group;
 OC *Bacillus*/Staphylococcus group; *Bacillus*.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RA MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,

RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus*
 halodurans and genomic sequence comparison with *Bacillus subtilis*."
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL: AP001511; BAB05019.1; -;
 DR InterPro: IPR000182; Acetyltransf_GCN5.
 DR Pfam: PF00583; Acetyltransf; 1.
 KW Complete proteome.
 SQ SEQUENCE 180 AA; 20375 MW; 100D7285E880BCEB CRC64;

Query Match 71.2%; Score 42; DB 16; Length 180;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ywgkgy 7
 |||||
 Db 102 YWGKGY 107

RESULT 14

O9KB15 PRELIMINARY; PRT; 181 AA.
 ID O9KB15;
 AC O9KB15;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE BH2121 PROTEIN.
 GN BH2121.
 OS *Bacillus halodurans*.
 OC Bacteria; Firmicutes; *Bacillus*/Clostridium group;
 OC *Bacillus*/Staphylococcus group; *Bacillus*.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus*
 halodurans and genomic sequence comparison with *Bacillus subtilis*."
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL: AP001514; BAB05840.1; -;
 DR InterPro: IPR000182; Acetyltransf_GCN5.
 DR Pfam: PF00583; Acetyltransf; 1.
 KW Complete proteome.
 SQ SEQUENCE 181 AA; 21429 MW; DB290C365101CA87 CRC64;

Query Match 71.2%; Score 42; DB 16; Length 181;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ywgkgy 7
 |||||
 Db 100 YWGKGY 105

RESULT 15

O9F189 PRELIMINARY; PRT; 446 AA.
 ID O9F189;
 AC O9F189;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE HYPOTHETICAL 47.8 KDA PROTEIN.
 OS *Alcaligenes eutrophus* (Ralstonia eutropha).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OX NCBI_TaxID=510;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RC STRAIN=335T;
RA Hinner I.-S., Buerger S., Schloemann M.;
RT "Characterization of a second gene cluster coding for enzymes of
RL catechol catabolism in Raistonia eutropha 335T.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF042280; AAC42023.1; -.
DR InterPro; IPR003818; SCFA_trans.
DR Pfam; PF02667; SCFA_trans; 1.
KW Hypothetical protein
SQ SEQUENCE 446 AA; 47785 MW; 7CDA57669C3CE0A5 CRC64;

Query Match          71.2%; Score 42; DB 2; Length 446;
Best Local Similarity 83.3%; Pred. No. 83;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 3 wgkyw 8
   |||||
Db 49 WKGFW 54

```

Search completed: August 12, 2002, 17:28:14
Job time: 361 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:18:53 ; Search time 21.42 seconds
(without alignments)
9.123 Million cell updates/sec

Title: 10-071247-1
Perfect score: 59
Sequence: 1 gywgkgyw 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	67.8	142	2	US-08-860-174A-7
2	40	67.8	274	2	US-08-860-174A-12
3	40	67.8	282	2	US-08-860-174A-10
4	39	66.1	624	3	US-08-947-965-78
5	39	66.1	655	1	US-08-469-202-27
6	39	66.1	655	1	US-08-469-202-28
7	39	66.1	655	2	US-08-484-434C-34
8	39	66.1	655	2	US-08-484-434C-35
9	38	64.4	392	1	US-08-423-441-2
10	38	64.4	1365	6	5194600-4
11	37	62.7	29	1	US-07-853-408B-99
12	37	62.7	29	2	US-08-308-865-99
13	37	62.7	29	4	US-09-042-353-296
14	37	62.7	29	4	US-08-758-417A-144
15	37	62.7	30	1	US-08-053-131-99
16	37	62.7	30	1	US-08-645-641-99
17	37	62.7	30	2	US-08-096-762-99
18	37	62.7	30	5	PCT-US92-10983-99
19	37	62.7	119	1	US-08-467-393-2
20	37	62.7	130	1	US-08-398-613A-22
21	37	62.7	130	1	US-08-398-612A-22
22	37	62.7	130	1	US-08-398-611A-22
23	37	62.7	130	2	US-08-491-334A-22
24	37	62.7	130	3	US-09-027-449-19
25	37	62.7	130	3	US-08-804-444A-19
26	37	62.7	130	4	US-09-026-985-19
27	37	62.7	251	1	US-08-398-612A-30

Sequence 30, Appl
Sequence 30, Appl
Sequence 27, Appl
Sequence 27, Appl
Sequence 27, Appl
Sequence 30, Appl
Sequence 4, Appl
Sequence 1, Appl
Sequence 12, Appl
Sequence 10, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 11, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 58, Appl

ALIGNMENTS

RESULT 1
US-08-860-174A-7
; Sequence 7, Application US/08860174A
; Patent No. 5989830
; GENERAL INFORMATION:
; APPLICANT: DAVIS, Paul James
; APPLICANT: VAN DER LOGT, Cornelis Paul Erik
; APPLICANT: VERHOEIJEN, Martine Elisa
; APPLICANT: WILSON, Steve
; TITLE OF INVENTION: A BIFUNCTIONAL OR BIVALENT
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESS: PILLSBURY, MADISON & SUTRO LLP
; STREET: 1100 New York Avenue, N.W.
; STREET: 9th Floor, East Tower
; CITY: WASHINGTON, D.C.
; STATE:
; COUNTRY: UNITED STATES
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS DOS Text
; SOFTWARE: MS WORD, WINDOWS 95, Ver. 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/860,174A
; FILING DATE: June 16, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95307332.7
; FILING DATE: October 16, 1995
; APPLICATION NUMBER: PCT/EP/96/03605
; FILING DATE: August 14, 1996
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 142 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-860-174A-7

Query Match 67.8%; Score 40; DB 2; Length 142;
Best Local Similarity 75.0%; Pred. No. 24;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 gywgkgyw 8
Db 122 GYKGKGYF 129

```
RESULT 2
US-08-860-174A-12
; Sequence 12, Application US/08860174A
; Patent No. 5989830
; GENERAL INFORMATION:
; APPLICANT: DAVIS, Paul James
; APPLICANT: VAN DER LOGT, Cornelis Paul Erik
; APPLICANT: VERHOEIJEN, Martine Elisa
; APPLICANT: WILSON, Steve
; TITLE OF INVENTION: A BIFUNCTIONAL OR BIVALENT
; TITLE OF INVENTION: ANTIBODY FRAGMENT ANALOGUE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY, MADISON & SUTRO LLP
; STREET: 1100 New York Avenue, N.W.
; STREET: 9th Floor, East Tower
; CITY: WASHINGTON, D.C.
; STATE:
; COUNTRY: UNITED STATES
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS DOS Text
; SOFTWARE: MS WORD, WINDOWS 95, Ver. 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/860,174A
; FILING DATE: June 16, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95307332.7
; FILING DATE: October 16, 1995
; APPLICATION NUMBER: PCT/EP/96/03605
; FILING DATE: August 14, 1996
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 274 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-860-174A-12
```

```
Query Match 67.8%; Score 40; DB 2; Length 274;
Best Local Similarity 75.0%; Pred. No. 44;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 gywgkgyw 8
||:||||:
Db 122 GYGKGYF 129
```

```
RESULT 3
US-08-860-174A-10
; Sequence 10, Application US/08860174A
; Patent No. 5989830
; GENERAL INFORMATION:
; APPLICANT: DAVIS, Paul James
; APPLICANT: VAN DER LOGT, Cornelis Paul Erik
; APPLICANT: VERHOEIJEN, Martine Elisa
; APPLICANT: WILSON, Steve
; TITLE OF INVENTION: A BIFUNCTIONAL OR BIVALENT
; TITLE OF INVENTION: ANTIBODY FRAGMENT ANALOGUE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY, MADISON & SUTRO LLP
; STREET: 1100 New York Avenue, N.W.
; STREET: 9th Floor, East Tower
; CITY: WASHINGTON, D.C.
; STATE:
; COUNTRY: UNITED STATES
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS DOS Text
; SOFTWARE: MS WORD, WINDOWS 95, Ver. 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/860,174A
; FILING DATE: June 16, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95307332.7
; FILING DATE: October 16, 1995
; APPLICATION NUMBER: PCT/EP/96/03605
; FILING DATE: August 14, 1996
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 282 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-860-174A-10
```

```
Query Match 67.8%; Score 40; DB 2; Length 282;
Best Local Similarity 75.0%; Pred. No. 45;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 gywgkgyw 8
||:||||:
Db 122 GYGKGYF 129
```

```
RESULT 4
US-08-947-965-78
; Sequence 78, Application US/08947965A
; Patent No. 6004790
; GENERAL INFORMATION:
; APPLICANT: Dijkhuizen, Lubbert
; APPLICANT: Dijkstra, Bauke
; APPLICANT: Andersen, Carsten
; APPLICANT: Osten, Claus von der
; TITLE OF INVENTION: Cyclomaltodextrin Glucanotransferase
; TITLE OF INVENTION: Variants
; FILE REFERENCE: 4285.204-US
; CURRENT APPLICATION NUMBER: US/08/947,965A
; CURRENT FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 0477/95
; EARLIER FILING DATE: 1995-04-21
; EARLIER APPLICATION NUMBER: 1173/95
; EARLIER FILING DATE: 1995-10-17
; EARLIER APPLICATION NUMBER: 1281/95
; EARLIER FILING DATE: 1995-11-16
; EARLIER APPLICATION NUMBER: PCT/DK96/00179
; EARLIER FILING DATE: 1996-04-22
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 78
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-08-947-965-78
```

```
Query Match 66.1%; Score 39; DB 3; Length 624;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 gywgkgyw 8
||:||||:
Db 91 GYGKGYF 98
```

```
RESULT 5
US-08-469-202-27
; Sequence 27, Application US/08469202
```

```
; Patent No. 5750875
; GENERAL INFORMATION:
; APPLICANT: STALKER, DAVID
; TITLE OF INVENTION: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CALGENE, INC.
; STREET: 1920 FIFTH STREET
; CITY: DAVIS
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1 (a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,202
; FILING DATE: 6-JUNE-95
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/016,881
; FILING DATE: 11 FEB 1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth Lassen
; REGISTRATION NUMBER: 31,845
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 93-2
; TELEPHONE: 916-753-6313
; TELEFAX: 916-753-1510
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 655 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-469-202-27

Query Match 66.1%; Score 39; DB 1; Length 655;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gywkggyw 8
Db 122 GYWGGRDYF 129

RESULT 6
US-08-469-202-28
; Sequence 28, Application US/08469202
; Patent No. 5750875
; GENERAL INFORMATION:
; APPLICANT: STALKER, DAVID
; TITLE OF INVENTION: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CALGENE, INC.
; STREET: 1920 FIFTH STREET
; CITY: DAVIS
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB

; Patent No. 5750875
; GENERAL INFORMATION:
; APPLICANT: STALKER, DAVID
; TITLE OF INVENTION: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CALGENE, INC.
; STREET: 1920 FIFTH STREET
; CITY: DAVIS
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
```

```
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1 (a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,202
; FILING DATE: 6-JUNE-95
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/016,881
; FILING DATE: 11 FEB 1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth Lassen
; REGISTRATION NUMBER: 31,845
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 93-2
; TELEPHONE: 916-753-6313
; TELEFAX: 916-753-1510
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 655 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-469-202-28

Query Match 66.1%; Score 39; DB 1; Length 655;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gywkggyw 8
Db 122 GYWGGRDYF 129

RESULT 7
US-08-484-434C-34
; Sequence 34, Application US/08484434C
; Patent No. 5969214
; GENERAL INFORMATION:
; APPLICANT: STALKER, DAVID
; TITLE OF INVENTION: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene LLC
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC
; OPERATING SYSTEM: Windows NT 4.0; MS Word for Windows 7.0a
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,434C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/016,881
; FILING DATE: 11-FEB-1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Schwedler, Carl J.
; REGISTRATION NUMBER: 36,924
```

REFERENCE/DOCKET NUMBER: 15593/01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 530-792-2265
TELEFAX: 530-792-2463
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 655 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-484-434C-34

Query Match 66.1%; Score 39; DB 2; Length 655;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 gywkggyw 8
| | | | |
Db 122 GYWRDIF 129

RESULT 8
US-08-484-434C-35
Sequence 35, Application US/08484434C
Patent No. 5969214
GENERAL INFORMATION:
APPLICANT: STALKER, DAVID
TITLE OF INVENTION: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene LLC
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC
OPERATING SYSTEM: Windows NT 4.0; MS Word for Windows 7.0a
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,434C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/016,881
FILING DATE: 11-FEB-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Schwedler, Carl J.
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: 15593/01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 530-792-2265
TELEFAX: 530-792-2463
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 655 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-484-434C-35

Query Match 66.1%; Score 39; DB 2; Length 655;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 gywkggyw 8
| | | | |
Db 122 GYWRDIF 129
RESULT 9
US-08-423-441-2
Sequence 2, Application US/08423441
Patent No. 5529926
GENERAL INFORMATION:
APPLICANT: MAAT, JAN
APPLICANT: MUSTERS, WOUTER
APPLICANT: STAM, HEIN
APPLICANT: SCHAAP, PETER J.
APPLICANT: VAN DE VONDERVOORT, PETER J.
APPLICANT: VISSER, JACOB
APPLICANT: VERBAKE, JOHANNES M.
TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA ENCODING A
TITLE OF INVENTION: RIPENING FORM OF A POLYPEPTIDE HAVING SULFHYDRYL OXIDASE
TITLE OF INVENTION: ACTIVITY
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/423,441
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/044,620
FILING DATE: 09-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16773
REFERENCE/DOCKET NUMBER: 202744/T7019(V)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 248453CUSH
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 392 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-423-441-2

Query Match 64.4%; Score 38; DB 1; Length 392;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 wkggyw 8
| | | | |
Db 154 WKGW 159

RESULT 10
5194600-4
Patent No. 5194600
APPLICANT: BUSSEY, HOWARD; BOONE, CHARLES; SOMMER, STEVE S.;
HILL, KATHRYN; MEADEN, PHILIP
TITLE OF INVENTION: GENES WHICH PARTICIPATE IN B-GLUCAN

ASSEMBLY AND USE THEREOF
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,316
FILING DATE: 05-MAR-1990
SEQ ID NO:4:
LENGTH: 1365
5194600-4

Query Match 64.4%; Score 38; DB 6; Length 1365;
Best Local Similarity 62.5%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gywkgkw 8
|||:|
Db 1234 GYWKIFYW 1241

RESULT 11

US-07-853-408B-99
Sequence 99, Application US/07853408B
Patent No. 5789650

GENERAL INFORMATION:
APPLICANT: Lonberg, Nils
APPLICANT: Kay, Robert M.
TITLE OF INVENTION: Transgenic No. 5789650-Human Animals for
Producing Heterologous Antibodies
NUMBER OF SEQUENCES: 150
CORRESPONDENCE ADDRESS:
ADDRESSEE: William M. Smith
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/853,408B
FILING DATE: 19920318
CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 14643-9
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 99:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-853-408B-99

Query Match 62.7%; Score 37; DB 1; Length 29;
Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gywkgkw 6
|||:|
Db 13 GYWGQG 18

RESULT 12

US-08-308-865-99

Sequence 99, Application US/08308865
Patent No. 5877397
GENERAL INFORMATION:
APPLICANT: Lonberg, Nils
APPLICANT: Kay, Robert M.
TITLE OF INVENTION: Transgenic No. 5877397-Human Animals for
Producing Heterologous Antibodies
NUMBER OF SEQUENCES: 150
CORRESPONDENCE ADDRESS:
ADDRESSEE: William M. Smith
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/308,865
FILING DATE:

CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/145,707
FILING DATE:
APPLICATION NUMBER: US 07/904,068
FILING DATE: 23-JUN-1992
ATTORNEY/AGENT INFORMATION:

NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 14643-9-1-1
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422

INFORMATION FOR SEQ ID NO: 99:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-308-865-99

Query Match 62.7%; Score 37; DB 2; Length 29;
Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gywkgkw 6
|||:|
Db 13 GYWGQG 18

RESULT 13

US-09-042-353-296
Sequence 99, Application US/09042353
Patent No. 6255458

GENERAL INFORMATION:
APPLICANT: Lonberg, Nils
APPLICANT: Kay, Robert M.
TITLE OF INVENTION: Transgenic No. 6255458-Human Animals for
Producing Heterologous Antibodies
NUMBER OF SEQUENCES: 421
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,353
FILING DATE:

CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/042,353
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Townsend and Crew LLP
REGISTRATION NUMBER: 421
REFERENCE/DOCKET NUMBER: 042-353-296
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 99:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-042-353-296

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/042,353
;; FILING DATE: 13-MAR-1998
;; CLASSIFICATION: 800
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/810,279
;; FILING DATE: 17-DEC-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/853,408
;; FILING DATE: 18-MAR-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/904,068
;; FILING DATE: 23-JUN-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/990,860
;; FILING DATE: 16-DEC-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/053,131
;; FILING DATE: 26-APR-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/096,762
;; FILING DATE: 22-JUL-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/155,301
;; FILING DATE: 18-NOV-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/161,739
;; FILING DATE: 03-DEC-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/165,699
;; FILING DATE: 10-DEC-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/209,741
;; FILING DATE: 09-MAR-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/352,322
;; FILING DATE: 07-DEC-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/544,404
;; FILING DATE: 10-OCT-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/728,463
;; FILING DATE: 10-OCT-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US96/16433
;; FILING DATE: 10-OCT-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/758,417
;; FILING DATE: 02-DEC-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US97/21803
;; FILING DATE: 01-DEC-1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 014643-009040US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 296:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 29 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-09-042-353-296

Query Match 62.7%; Score 37; DB 4; Length 29;
Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 gywgkg 6
Db 13 GYGQG 18
||||:|

RESULT 14
US-08-758-417A-144
; Sequence 144, Application US/08/758417A
; Patent No. 6300129
; GENERAL INFORMATION:
; APPLICANT: Lonberg, Nils
; Kay, Robert M.
; TITLE OF INVENTION: Transgenic No. 6300129-Human Animals for
; Producing Heterologous Antibodies
; NUMBER OF SEQUENCES: 417
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/758.417A
; FILING DATE: 02-DEC-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/728,463
; FILING DATE: 10-OCT-1996
; APPLICATION NUMBER: US 08/544,404
; FILING DATE: 10-OCT-1995
; APPLICATION NUMBER: US 08/352,322
; FILING DATE: 07-DEC-1994
; APPLICATION NUMBER: US 08/209,741
; FILING DATE: 09-MAR-1994
; APPLICATION NUMBER: US 08/165,699
; FILING DATE: 10-DEC-1993
; APPLICATION NUMBER: US 08/161,739
; FILING DATE: 03-DEC-1993
; APPLICATION NUMBER: US 08/155,301
; FILING DATE: 18-NOV-1993
; APPLICATION NUMBER: US 08/096,762
; FILING DATE: 22-JUL-1993
; APPLICATION NUMBER: US 08/053,131
; FILING DATE: 26-APR-1993
; APPLICATION NUMBER: US 07/990,860
; FILING DATE: 16-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Serafini, Andrew T.
; REGISTRATION NUMBER: 41,303
; REFERENCE/DOCKET NUMBER: 014643-009030US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 144:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 144:
US-08-758-417A-144

Job time: 197 sec

Query Match 62.7%; Score 37; DB 4; Length 29;
Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 gywqkg 6
|||||
Db 13 GYWGQ 18

RESULT 15
US-08-053-131-99
; Sequence 99, Application US/08053131
; Patent No. 5661016
; GENERAL INFORMATION:
; APPLICANT: Lonberg, Nils
; APPLICANT: Kay, Robert M.
; TITLE OF INVENTION: Transgenic No. 5661016-Human Animals for
; TITLE OF INVENTION: Producing Heterologous Antibodies
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 200
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/053,131
; FILING DATE: 26-APR-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/990,860
; FILING DATE: 16-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/810,279
; FILING DATE: 17-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/853,408
; FILING DATE: 18-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 14643-9-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-053-131-99

Query Match 62.7%; Score 37; DB 1; Length 30;
Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 gywqkg 6
|||||
Db 14 GYWGQ 19

Search completed: August 12, 2002, 17:22:10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:22:44 ; Search time 27.08 Seconds
(without alignments)
28.387 Million cell updates/sec

Title: 10-071247-2
Perfect score: 66
Sequence: 1 cywgcgyw 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	46	69.7	554	1 A28716	saposin precursor
2	45.5	68.9	269	2 E96704	hypothetical prote
3	45.5	68.9	283	2 H84430	hypothetical prote
4	42	63.6	48	2 A49739	relaxin - horse (f
5	42	63.6	143	2 T47053	relaxin B,C and A
6	42	63.6	342	2 T31757	hypothetical prote
7	42	63.6	347	2 T31755	hypothetical prote
8	42	63.6	351	2 T31758	hypothetical prote
9	42	63.6	473	2 I54210	N-acetylglactosam
10	42	63.6	533	1 KJH0AB	N-acetylglactosam
11	42	63.6	863	2 D70770	probable glycogen
12	41	62.1	83	2 B72392	hypothetical prote
13	41	62.1	280	2 G84839	late embryogenesis
14	41	62.1	441	2 H96968	integral membrane
15	41	62.1	589	2 B29514	muscarinic acetylch
16	41	62.1	589	2 A29476	muscarinic acetylch
17	41	62.1	590	2 S01114	muscarinic acetylch
18	41	62.1	590	2 S47572	muscarinic acetylch
19	41	62.1	590	2 AF1810	hypothetical prote
20	41	62.1	741	2 S73827	hypothetical prote
21	41	62.1	2970	2 T08839	polyprotein - marm
22	40	60.6	113	2 Pti063	fg heavy chain v r
23	40	60.6	130	1 A69894	hypothetical prote
24	40	60.6	130	2 T44809	hypothetical prote
25	40	60.6	396	2 T04561	hypothetical prote
26	40	60.6	415	2 H86204	probable phosphoes
27	40	60.6	441	2 F83694	short-chain fatty
28	40	60.6	456	2 T40367	n-acetylglucosamin
29	40	60.6	499	2 T36462	hypothetical prote

ALIGNMENTS

RESULT 1
A28716
saposin precursor - rat
N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; compon
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulf
N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A28716
R:Collard, M.W.; Sylvester, S.R.; Tsuruta, J.K.; Griswold, M.D.
Biochemistry 27, 4557-4564, 1998
A:Title: Biosynthesis and molecular cloning of sulfated glycoprotein 1 secreted by ra
A:Reference number: A28716; MUID:89000647
A:Accession: A28716
A:Molecule type: mRNA
A:Residues: 1-554 <COL>
A:Cross-references: GB:M19936; NID:g206904; PIDN:AAA2136.1; PID:g206905
A:Note: parts of this sequence, including the amino end of the mature protein, were d
C:Function:
A:Description: saposins bind sphingolipids, form hydrophilic complexes and make them
A:Pathway: sphingolipid catabolism
A:Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-gluc
A:Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by aryls
A:Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiast
C:Superfamily: saposin; saposin repeat homology
C:Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome;
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-554/Product: prosaposin #status predicted <PRO>
F:55-148/Domain: saposin repeat homology <SAP1>
F:60-143/Product: saposin A #status predicted <SAPA>
F:189-280/Domain: saposin repeat homology <SAP2>
F:194-273/Product: saposin B #status predicted <SAB1>
F:306-387/Domain: saposin repeat homology <SAP3>
F:310-389/Product: saposin C #status predicted <SAPC>
F:431-522/Domain: saposin repeat homology <SAP4>
F:437-514/Product: saposin D #status predicted <SAPD>
F:63-138, 66-132, 94-106, 439-512, 442-506, 470-481/Disulfide bonds: #status predicted
F:80, 214, 331, 456/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:197-270, 200, 260-264, 229-240, 314-387, 317-381, 345-356/Disulfide bonds: #status predicted

Query Match 69.7%; Score 46; DB 1; Length 554;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 cywgcgyw 8
| | | | |
Db 525 CWMGPGYW 532

RESULT 2
E96704

hypothetical protein T23K23.3 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: E96704
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Mafti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: E96704
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-269 <STO>
A:Cross-references: GB:AE005173; NID:g6553884; PIDN:AAF16550.1; GSPDB:GN00141
C:Genetics:
A:Gene: T23K23.3
A:Map position: 1

Query Match 68.9%; Score 45.5; DB 2; Length 269;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 cy-wgcyyw 8
|| |||||
Db 190 CYRWGCGGW 198

RESULT 3
H84430
hypothetical protein At2g01930 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: H84430
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moifatt, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: H84430
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-283 <STO>
A:Cross-references: GB:AE002093; NID:g6598329; PIDN:AAF18588.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g01930
A:Map position: 2

Query Match 68.9%; Score 45.5; DB 2; Length 283;
Best Local Similarity 77.8%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 cy-wgcyyw 8
|| |||||
Db 204 CYRWGCGGW 212

RESULT 4
A49739
relaxin - horse (fragments)
C:Species: Equus caballus (domestic horse)
C>Date: 07-Apr-1994 #sequence_revision 26-May-1995 #text_change 26-May-1995
C:Accession: B49739; A49739
R:Stewart, D.R.; Nevins, B.; Hadas, E.; Vandlen, R.
Endocrinology 129, 375-383, 1991

A:Title: Affinity purification and sequence determination of equine relaxin.
A:Reference number: A49739; MUID:91275796
A:Accession: B49739
A:Molecule type: protein
A:Residues: 1-28 <STE>
A:Accession: A49739
A:Molecule type: protein
A:Residues: 29-48 <ST2>
C:Superfamily: insulin
C:Keywords: hormone; pyroglutamic acid
F;1-28/29-48/Product: relaxin #status experimental <MAT>
F;1-28/Domain: chain B #status experimental <CHB>
F;29-48/Domain: chain A #status experimental <CHA>
F;1.29/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status expe
F;10-34,22-39,35-48/Disulfide bonds: #status predicted

Query Match 63.6%; Score 42; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cywgc 5
|||||
Db 35 CYWGC 39

RESULT 5
I47053
relaxin B.C and A chains - horse (fragment)
C:Species: Equus sp
C>Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 16-Jul-1999
C:Accession: I47053
R:Klonisch, T.; Ryan, P.L.; Yamashiro, S.; Porter, D.G.
Biol. Reprod. 52, 1307-1315, 1995
A:Title: Partial complementary deoxyribonucleic acid cloning of equine relaxin messen
A:Reference number: I47053; MUID:95359320
A:Accession: I47053
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-143 <KLO>
A:Cross-references: GB:S78800; NID:g1042059; PIDN:AAB35036.1; PID:g1042060
C:Superfamily: insulin

Query Match 63.6%; Score 42; DB 2; Length 143;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cywgc 5
|||||
Db 138 CYWGC 142

RESULT 6
T31757
hypothetical protein C07G3.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T31757
R:Geisler, C.; Wamsley, P.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid C07G3.
A:Reference number: Z21080
A:Accession: T31757
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-342 <GEI>
A:Cross-references: EMBL:AF016432; PIDN:AAB65382.1; GSPDB:GN00023; CESP:C07G3.4
A:Experimental source: strain Bristol N2; clone C07G3
C:Genetics:
A:Gene: CESP:C07G3.4
A:Map position: 5
A:Introns: 66/3; 122/2; 167/2; 214/3; 319/1

Query Match 63.6%; Score 42; DB 2; Length 342;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cywgc 5
|
|
|
|
Db 97 CYWGC 101

RESULT 7

T31755

hypothetical protein C07G3.6 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T31755

R:Geisel, C.; Wansley, P.

Submitted to the EMBL Data Library, July 1997

A:Description: The sequence of *C. elegans* cosmid C07G3.

A:Reference number: Z21080

A:Accession: T31755

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-347 <GEI>

A:Cross-references: EMBL:AF016432; PIDN:AAB65378.1; GSPDB:GN00023; CESP:C07G3.6

A:Experimental source: strain Bristol N2; clone C07G3

C:Genetics:

A:Gene: CESP:C07G3.6

A:Map position: 5

A:Introns: 66/3; 122/2; 167/2; 214/3; 319/1

Query Match

63.6%; Score 42; DB 2; Length 347;

Best Local Similarity 100.0%; Pred. No. 44;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cywgc 5

|

|

|

|

Db 97 CYWGC 101

RESULT 8

T31758

hypothetical protein C07G3.3 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T31758

R:Geisel, C.; Wansley, P.

Submitted to the EMBL Data Library, July 1997

A:Description: The sequence of *C. elegans* cosmid C07G3.

A:Reference number: Z21080

A:Accession: T31758

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-351 <GEI>

A:Cross-references: EMBL:AF016432; PIDN:AAB65377.1; GSPDB:GN00023; CESP:C07G3.3

A:Experimental source: strain Bristol N2; clone C07G3

C:Genetics:

A:Gene: CESP:C07G3.3

A:Map position: 5

A:Introns: 66/3; 122/2; 167/2; 214/3; 319/1

Query Match

63.6%; Score 42; DB 2; Length 351;

Best Local Similarity 100.0%; Pred. No. 45;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cywgc 5

|

|

|

|

Db 97 CYWGC 101

RESULT 9

I54210

N-acetylglactosamine-4-sulfatase (EC 3.1.6.12) precursor - rat (fragment)

N:Alternate names: arylsulfatase (EC 3.1.6.1) B [misidentification]

C:Species: *Rattus norvegicus* (Norway rat)

C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 02-Jun-2000

C:Accession: I54210

R:Kunieda, T.

Genomics 29, 582-587, 1995

A:Title: Mucopolysaccharidosis type VI in rats: Isolation of cDNAs encoding arylsulfatase

A:Reference number: I54210; MUID:96121368

A:Accession: I54210

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-473 <RES>

A:Cross-references: GB:D49434; NID:gl065603; PIDN:BAA08412.1; PID:g1089794

C:Genetics:

A:Gene: ARSB

C:Superfamily: animal sulfatase

C:Keywords: sulfuric ester hydrolase

F:31/Modified site: 3-oxoalanine (Cys) #status predicted

Query Match

63.6%; Score 42; DB 2; Length 473;

Best Local Similarity 85.7%; Pred. No. 56;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ywgcgyw 8

|

|

|

|

Db 384 YPGCGYW 390

RESULT 10

KJHUBB

N-acetylglactosamine-4-sulfatase (EC 3.1.6.12) precursor [validated] - human

N:Alternate names: arylsulfatase B (ASB); chondroitinase; chondroitinsulfatase; G4S;

C:Species: *Homo sapiens* (man)

C:Date: 31-Dec-1993 #sequence_revision 27-Oct-1995 #text_change 08-Dec-2000

C:Accession: S35990; S33307; A35078; A45659; A42449; B42449; C42449; I54217; A56865

R:Peters, C.W.B.

submitted to the EMBL Data Library, March 1993

A:Reference number: S35990

A:Accession: S35990

A:Molecule type: DNA

A:Residues: 1-533 <PEP>

A:Cross-references: EMBL:X72735; NID:g289009; PIDN:CAA51272.1; PID:g825628

R:Modaresi, S.; Rupp, K.; von Figura, K.; Peters, C.

Biol. Chem. Hoppe-Seyler 374, 327-335, 1993

A:Title: Structure of the human arylsulfatase B gene.

A:Reference number: S33307; MUID:93332648

A:Accession: S33307

A:Molecule type: DNA

A:Residues: 1-104 <MOD>

A:Cross-references: EMBL:X72736; EMBL:X72737; EMBL:X72738; EMBL:X72739;

A:Note: the enzyme is referred to as EC 3.1.6.9

R:Peters, C.; Schmidt, B.; Rommerskirch, W.; Rupp, K.; Zuehlendorf, M.; Vingron, M.; M

J. Biol. Chem. 265, 3374-3381, 1990

A:Title: Phylogenetic conservation of arylsulfatases. cDNA cloning and expression of

A:Reference number: A35078; MUID:90153994

A:Accession: A35078

A:Molecule type: mRNA

A:Residues: 1-357, 'V', 359-533 <PE2>

A:Cross-references: GB:J05225; NID:g179076; PIDN:AAA51784.1; PID:g179077

A:Note: parts of this sequence were determined by protein sequencing

R:Litjens, T.; Morris, C.P.; Gibson, G.J.; Beckmann, K.R.; Hopwood, J.J.

Biochem. Int. 24, 209-215, 1991

A:Title: Human N-acetylglactosamine-4-sulphatase: protein maturation and isolation o

A:Reference number: A45659; MUID:92028992

A:Accession: A45659

A:Molecule type: DNA; protein

A:Residues: 1-104 <LIT>

A:Cross-references: GB:S57777; NID:g236697; PIDN:AAB19988.1; PID:g236698

A:Note: sequence extracted from NCBI backbone (NCBIN:57777, NCBIP:57778)
A:Note: the enzyme is referred to as EC 3.1.6.1
A:Note: parts of this sequence, including the amino end of the mature protein, were determined from the mature protein
A:Note: a form is described with a proteolytic cleavage somewhere between residue 450 and 451
R:Jin, W.D.; Jackson, C.E.; Desnick, R.J.; Schuchman, E.H.
Am. J. Hum. Genet. 50, 795-800, 1992
A:Title: Mucopolysaccharidosis type VI: identification of three mutations in the arylsulphatase gene
A:Reference number: A42449; MUID:92197625
A:Accession: A42449
A:Molecule type: mRNA
A:Residues: 115-116, 'R', 118 <JIN>
A:Cross-references: GB:S90729; NID:g247486; PIDN:AA21833.1; PID:g247487
A:Note: sequence extracted from NCBI backbone (NCBIN:90729, NCBIP:90731)
A:Accession: B42449
A:Molecule type: mRNA
A:Residues: 234-235, 'P', 237-238 <JIT2>
A:Cross-references: GB:S90736; NID:g247488; PIDN:AA21832.1; PID:g247489
A:Note: sequence extracted from NCBI backbone (NCBIN:90736, NCBIP:90739)
A:Accession: C42449
A:Molecule type: mRNA
A:Residues: 403-404, 'Y', 406-407 <JIT3>
A:Cross-references: GB:S90743; NID:g247490; PIDN:AA21833.1; PID:g247491
A:Note: sequence extracted from NCBI backbone (NCBIN:90743, NCBIP:90747)
A:Note: the enzyme is referred to as EC 3.1.6.1
A:Note: these mutations give rise to mucopolysaccharidosis type VI, Maroteaux-Lamy disease
R:Schmidt, B.; Selmer, T.; Ingendoh, A.; von Figura, K.
Cell 82, 271-278, 1995
A:Title: A novel amino acid modification in sulfatases that is defective in multiple sulfatase deficiency
A:Reference number: A57113; MUID:95354208
A:Contents: annotation; identification of 3-oxoalanine, 2-amino-3-oxopropanoic acid
R:Schuchman, E.H.; Jackson, C.E.; Desnick, R.J.
Genomics 6, 149-158, 1990
A:Title: Human arylsulfatase B: MOPAC cloning, nucleotide sequence of a full-length cDNA
A:Reference number: I54217; MUID:90152677
A:Accession: I54217
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-357, 'V', 359-375, 'M', 377-533 <RES>
A:Cross-references: GB:M32373; NID:gl79029; PIDN:AAA51779.1; PID:gl79030
R:Kobayashi, T.; Honke, K.; Jin, T.; Gasa, S.; Miyazaki, T.; Makita, A.
Biochim. Biophys. Acta 1159, 243-247, 1992
A:Title: Components and proteolytic processing sites of arylsulfatase B from human placenta
A:Reference number: A56865; MUID:93003385
A:Accession: A56865
A:Molecule type: protein
A:Residues: 41-55; 424-425, 'X', 427-454; 466-483 <KOB>
A:Experimental source: placenta
A:Note: sequence modified after extraction from NCBI backbone
A:Note: the fragments shown are the amino ends of the alpha, gamma, and beta chains of the enzyme
C:Comment: This enzyme is frequently misidentified as EC 3.1.6.1.
C:Genetics:
A:Gene: GDB:ARSB
A:Cross-references: GDB:119008; OMIM:253200
A:Map position: 5q11-q13
A:Introns: 104/3; 167/1; 230/3; 300/1; 381/2; 405/1; 446/1
A:Note: defects in this gene can cause mucopolysaccharidosis type VI, Maroteaux-Lamy disease
C:Function:
A:Description: hydrolyzes N-acetylgalactosamine-4-sulfate units in chondroitin sulfate A and B
C:Superfamily: animal sulfatase
C:Keywords: glycoprotein; lysosomal storage disease; lysosome; sulfuric ester hydrolase
F:1-423/Product: signal sequence #status predicted <SIG>
F:41-423/Product: alpha chain #status predicted <AMAT>
F:424-465/Product: gamma chain #status predicted <GMAT>
F:466-533/Product: beta chain #status experimental <BMAT>
F:91/Modified site: 3-oxoalanine (Cys) #status experimental
F:188, 279, 366, 458/Binding site: carboxylate (Asn) (covalent) #status predicted
F:291/Binding site: carbohydrate (Asn) (covalent) #status absent
F:426/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 63.6%; Score 42; DB 1; Length 533;
Best Local Similarity 85.7%; Pred. No. 62;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 ywgcgyw 8
| | | | |
DB 444 YPGCGYW 450

RESULT 11

D70770
probable glycogen phosphorylase - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: D70770
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Feltwell, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Spares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: D70770
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-863 <COL>
A:Cross-references: GB:Z73902; GB:AL123456; NID:g3261576; PIDN:CAA98092.1; PID:e24502
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: glgP

Query Match 63.6%; Score 42; DB 2; Length 863;
Best Local Similarity 71.4%; Pred. No. 89;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ywgcgyw 8
| | | | |
DB 595 YWGCDDW 601

RESULT 12

B72392
hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: B72392
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A:Reference number: A72200; MUID:99287316
A:Accession: B72392
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-83 <ARN>
A:Cross-references: GB:AE001713; GB:AE000512; NID:g4980809; PIDN:AAD35403.1; PID:g498
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0315

Query Match 62.1%; Score 41; DB 2; Length 83;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 cywgcgyw 8
| | | | |
DB 5 CFWSWGF 12

RESULT 13

G84839
late embryogenesis abundant M17 protein [Imported] - Arabidopsis thaliana

```
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: G84839
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487
A;Accession: G84839
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-280 <STO>
A;Cross-references: GB:AE002093; NID:g3894196; PIDN:AAK78545.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g41260
A;Map position: 2

Query Match 62.1%; Score 41; DB 2; Length 280;
Best Local Similarity 60.0%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

QY 1 cywgc--gyw 8
DB 78 CRWCCGGW 87

RESULT 14
H96968
Integral membrane protein similar to antibiotic resistance protein B. subtilis [imported]
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: H96968
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: H96968
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-441 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK78539.1; PID:gi5023427; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC0560

Query Match 62.1%; Score 41; DB 2; Length 441;
Best Local Similarity 71.4%; Pred. No. 74;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 cywgcgy 7
DB 333 CFWMGM 339

RESULT 15
B29514
muscarinic acetylcholine receptor M3 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 18-Feb-2000
C;Accession: B94518; B94293; B37121; B29514
R;Bonner, T.I.
submitted to GenBank, July 1987
A;Reference number: A94518
A;Accession: B94518
A;Molecule type: mRNA
A;Residues: 1-589 <BOI>
R;Bonner, T.I.; Buckley, N.J.; Young, A.C.; Brann, M.R.
Science 237, 527-532, 1987
A;Title: Identification of a family of muscarinic acetylcholine receptor genes.
```

```
A;Reference number: A94293; MUID:87263421
A;Accession: B94293
A;Molecule type: mRNA
A;Residues: 1-269;463-589 <BO2>
A;Experimental source: cerebral cortex
A;Note: only a part of the protein translation is given; none of the nucleotide sequen
R;Kutenbach, E.; Curtis, C.A.M.; Pedder, E.K.; Aitken, A.; Harris, A.C.M.; Hulme, E.
J. Biol. Chem. 265, 13702-13708, 1990
A;Title: Muscarinic acetylcholine receptors. Peptide sequencing identifies residues i
A;Reference number: A37121; MUID:90337982
A;Accession: B37121
A;Status: preliminary
A;Molecule type: protein
A;Residues: 104-166 <KUR>
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; tran
F;67-90/Domain: transmembrane #status predicted <TM1>
F;104-124/Domain: transmembrane #status predicted <TM2>
F;142-163/Domain: transmembrane #status predicted <TM3>
F;184-206/Domain: transmembrane #status predicted <TM4>
F;230-251/Domain: transmembrane #status predicted <TM5>
F;492-512/Domain: transmembrane #status predicted <TM6>
F;527-545/Domain: transmembrane #status predicted <TM7>
F;6,15,41,48,52/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 62.1%; Score 41; DB 2; Length 589;
Best Local Similarity 71.4%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ywgcgyw 8
DB 524 YWNLGVW 530

Search completed: August 12, 2002, 17:22:45
Job time: 182 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:27:25 ; Search time 13.51 seconds
(without alignments)
22.928 Million cell updates/sec

Title: 10-071247-2
Perfect score: 66
Sequence: 1 cywgcgyw 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	69.7	518	1 SAP_CHICK	O13035 gallus gall
2	46	69.7	554	1 SAP_RAT	P10960 rattus norv
3	42	63.6	182	1 RELX_HORSE	P22969 equus cabal
4	42	63.6	473	1 ARSB_RAT	P50430 rattus norv
5	42	63.6	533	1 ARSB_HUMAN	P15848 homo sapien
6	42	63.6	863	1 PHSG_MYCTU	Q10639 mycobacteri
7	41	62.1	589	1 ACN3_MOUSE	O9er23 mus musculu
8	41	62.1	589	1 ACN3_RAT	P04883 rattus norv
9	41	62.1	590	1 ACN3_BOVIN	P41984 bos tauru
10	41	62.1	590	1 ACN3_PIG	P11483 sus scrofa
11	41	62.1	741	1 YD35_MYCPN	P75443 mycoplasma
12	40	60.6	524	1 SAP_HUMAN	P07602 h proactiva
13	40	60.6	525	1 SAP_BOVIN	P26779 b proactiva
14	40	60.6	557	1 SAP_MOUSE	O61207 mus musculu
15	40	60.6	573	1 INA1_TRIHA	P34054 trichoderma
16	39.5	59.8	61	1 K62_SHEEP	P02448 ovis aries
17	39	59.1	296	1 NIFD_NOSCO	P52337 nostoc comm
18	39	59.1	396	1 INPP_MOUSE	P49442 mus musculu
19	39	59.1	400	1 INPP_HUMAN	P49441 homo sapien
20	39	59.1	400	1 INPP_BOVIN	P21327 bos tauru
21	39	59.1	452	1 TRG_PLAFO	P34787 plasmodium
22	39	59.1	497	1 NIFD_ANASP	P00464 anabaena sp
23	38.5	58.3	665	1 ENV_MLVMO	P03385 moloney mur
24	38	57.6	105	1 RSNB_MOUSE	O99p86 mus musculu
25	38	57.6	105	1 RSNB_HUMAN	O9bq08 homo sapien
26	38	57.6	485	1 TRE2_SYNY3	P74130 synechocyst
27	37	56.1	79	1 KRA_RABIT	Q02957 oryctolagus
28	37	56.1	94	1 V093_FOWPV	Q9j5c7 rowipox vir
29	37	56.1	590	1 ACN3_GORGO	P20309 homo sapien
30	37	56.1	590	1 ACN3_HUMAN	Q9n2a3 gorilla gor
31	37	56.1	590	1 ACN3_PANTR	Q9n2a4 pan troglod
32	37	56.1	590	1 ACN3_PONPY	Q9n2a2 pongo pygma
33	37	56.1	1353	1 PUR2_DROME	P00967 d trifuncti

34	36.5	55.3	236	1 IPT_AGRVI	Q04590 agrobacteri
35	36	54.5	82	1 KRHA_SHEEP	Q02958 ovis aries
36	36	54.5	110	1 IBP_CARCR	P00993 caretta car
37	36	54.5	263	1 T2T8_THETH	P29748 thermus aqu
38	36	54.5	263	1 T2TA_THEAQ	P14386 thermus aqu
39	36	54.5	622	1 AMT3_CABEL	Q21565 caenorhabdi
40	36	54.5	1365	1 KRE5_YEAST	P22023 saccharomyc
41	36	54.5	1593	1 ATL2_HUMAN	P38397 homo sapien
42	35.5	53.8	433	1 YBL4_YEAST	P38211 saccharomyc
43	35.5	53.8	491	1 YIJ0_YEAST	P40499 saccharomyc
44	35	53.0	27	1 TXA3_ANESU	P01535 anemonia su
45	35	53.0	273	1 YFCO_ECOLI	P76498 escherichia

ALIGNMENTS

RESULT	1
SAP_CHICK	
ID	SAP_CHICK
AC	OL3035; STANDARD; PRT; 518 AA.
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DE	Proactivator polypeptide precursor [Contains: Saposin B;
DE	Saposin C; Saposin D].
GN	PSAP.
OS	Gallus gallus (Chicken).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC	Gallus.
OX	NCBI_TaxID=9031;
RN	[1]
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 194-203.
RC	TISSUE=Brain, and Liver;
RX	MEDLINE=98129745; PubMed=9461526;
RA	Azuma N., Seo H.-C., Lie O., Fu Q., Gould R.M., Hiraiwa M., Burt D.W.,
RA	Paton I.R., Morrice D.R., O'Brien J.S., Kishimoto Y.;
RT	"Cloning, expression and map assignment of chicken prosaposin.";
RL	Biochem. J. 330:321-327(1998).
RN	[2]
RP	SEQUENCE FROM N.A.
RA	Altman N., Horowitz M.;
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC	-!- FUNCTION: THE LYSOSOMAL DEGRADATION OF SPHINGOLIPIDS TAKES PLACE
CC	BY THE SEQUENTIAL ACTION OF SPECIFIC HYDROLASES. SOME OF THESE
CC	ENZYMES REQUIRE SPECIFIC LOW-MOLECULAR MASS, NON-ENZYMIC PROTEINS:
CC	THE SPHINGOLIPIDS ACTIVATOR PROTEINS (COPROTEINS) (BY SIMILARITY).
CC	-!- FUNCTION: SAPOSIN A AND SAPOSIN C STIMULATE THE HYDROLYSIS OF
CC	GLUCOSYLKERAMIDE BY BETA-GLUCOSYLKERAMIDASE (EC 3.2.1.45) AND
CC	GALACTOSYLKERAMIDE BY BETA-GALACTOSYLKERAMIDASE (EC 3.2.1.46).
CC	SAPOSIN-C APPARENTLY ACTS BY COMBINING WITH THE ENZYME AND ACIDIC
CC	LIPID TO FORM AN ACTIVATED COMPLEX, RATHER THAN BY SOLUBILIZING
CC	THE SUBSTRATE (BY SIMILARITY).
CC	-!- FUNCTION: SAPOSIN B STIMULATES THE HYDROLYSIS OF GALACTO-
CC	CEREROSIDE SULFATE BY ARYLSULFATASE A (EC 3.1.6.8), GM1
CC	GANGLIOSIDES BY BETA-GALACTOSIDASE (EC 3.2.1.23) AND
CC	GLOBOTRIOSYLKERAMIDE BY ALPHA-GALACTOSIDASE A (EC 3.2.1.22).
CC	SAPOSIN-B FORMS A SOLUBILIZING COMPLEX WITH THE SUBSTRATES OF THE
CC	SPHINGOLIPID HYDROLASES (BY SIMILARITY).
CC	-!- FUNCTION: SAPOSIN D IS A SPECIFIC SPHINGOMYELIN PHOSPHODIESTERASE
CC	ACTIVATOR (EC 3.1.4.12) (BY SIMILARITY).
CC	-!- SUBUNIT: SAPOSIN B IS A HOMODIMER (BY SIMILARITY).
CC	-!- SUBCELLULAR LOCATION: Lysosomal (By similarity).
CC	PEPTIDES, WHICH ARE SIMILAR TO EACH OTHER AND ARE SPHINGOLIPID
CC	HYDROLASE ACTIVATOR PROTEINS (BY SIMILARITY).
CC	-!- SIMILARITY: CONTAINS 2 SAPOSIN A-TYPE DOMAINS.
CC	-!- SIMILARITY: CONTAINS 4 SAPOSIN B-TYPE DOMAINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

TISSUE-Sertoli cells;
MEDLINE=89000647; PubMed=3048385;
Collard M.W., Sylvester S.R., Tsuruta J.K., Griswold M.D.;
"Biosynthesis and molecular cloning of sulfated glycoprotein 1
secreted by rat Sertoli cells: sequence similarity with the
70-kilodalton precursor to sulfatide/GM1 activator.";
Biochemistry 27:4557-4564(1988).
[2]
SEQUENCE FROM N.A.
RP
TISSUE-Testis;
MEDLINE=96128541; PubMed=8573994;
Morales C.R., El-Alfy M., Zhao Q., Igoudra S.A.;
"Molecular role of sulfated glycoprotein-1 (SGP-1/prosaposin) in
Sertoli cells.";
Histol. Histopathol. 10:1023-1034(1995).
[3]
SEQUENCE FROM N.A.
RP
TISSUE-Testicle;
MEDLINE=96175245; PubMed=8601692;
Morales C.R., El-Alfy M., Zhao Q., Igoudra S.A.;
"Expression and tissue distribution of rat sulfated glycoprotein-1
(prosaposin).";
J. Histochem. Cytochem. 44:327-337(1996).
CC
-|- SUBCELLULAR LOCATION: Extracellular.
CC
-|- SIMILARITY: CONTAINS 2 SAPOSIN A-TYPE DOMAINS.
CC
-|- SIMILARITY: CONTAINS 4 SAPOSIN B-TYPE DOMAINS.
CC
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AB003471; BAA19914.1; -
EMBL; AF108656; AAF05899.1; -
DR InterPro; IPR003119; SapA.
DR InterPro; IPR000004; SapB.
DR Pfam; PF02199; SAPA; 2.
DR ProDom; PD001732; SapB; 3.
DR ProDom; PD012321; Saposin; 1.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPB; 4.
KW Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat;
KW GM2-gangliosidosis.
FT SIGNAL 1
FT CHAIN 61 143 POTENTIAL.
FT CHAIN 194 276 SAPOSIN A.
FT CHAIN 307 387 SAPOSIN B.
FT CHAIN 399 480 SAPOSIN C.
FT CHAIN 422 55 SAPOSIN D.
FT DOMAIN 60 143 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 193 277 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 307 388 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 399 480 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 485 518 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 518 518 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 64 139 BY SIMILARITY.
FT DISULFID 67 133 BY SIMILARITY.
FT DISULFID 95 107 BY SIMILARITY.
FT DISULFID 197 273 BY SIMILARITY.
FT DISULFID 200 267 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 311 384 BY SIMILARITY.
FT DISULFID 314 378 BY SIMILARITY.
FT DISULFID 342 353 BY SIMILARITY.
FT DISULFID 403 476 BY SIMILARITY.
FT DISULFID 406 470 BY SIMILARITY.
FT DISULFID 434 445 BY SIMILARITY.
FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 94 94 R -> T (IN REF. 2).
FT CONFLICT 486 486 E -> D (IN REF. 2).
FT SEQUENCE 518 AA; 57601 MW; B803000E891C3963 CRC64;

Query Match 69.7%; Score 46; DB 1; Length 518;
Best Local Similarity 75.0%; Pred. No. 4.2;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CVWGCGYW 8
Db 489 CVWGPY 496

RESULT 2
SAP_RAT STANDARD; PRT; 554 AA.
AC P10960; Q62841; Q64190;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
GN PSAP OR SGP1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

EMBL; M19936; AAA42136.1; -
EMBL; S81353; AAB36042.2; -
EMBL; S81373; AAB36233.2; -
DR PIR; A28716; A28716.
DR InterPro; IPR003119; SapA.
DR InterPro; IPR000004; SapB.
DR Pfam; PF02199; SAPA; 2.
DR ProDom; PD001732; SapB; 3.
DR ProDom; PD012321; Saposin; 1.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPB; 4.
DR Sulfation; Signal; Glycoprotein; Repeat.
FT SIGNAL 1 16
FT CHAIN 17 554 SULFATED GLYCOPROTEIN 1.
FT DOMAIN 21 54 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 61 138 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 193 274 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 310 391 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 435 516 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 521 554 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 63 138 BY SIMILARITY.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 197 270 BY SIMILARITY.
FT DISULFID 200 264 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 314 387 BY SIMILARITY.
FT DISULFID 317 381 BY SIMILARITY.
FT DISULFID 345 356 BY SIMILARITY.
FT DISULFID 439 512 BY SIMILARITY.
FT DISULFID 442 506 BY SIMILARITY.
FT DISULFID 470 481 BY SIMILARITY.
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 115 115 P -> L (IN REF. 2 AND 3).

FT CONFLICT 299 299 D -> E (IN REF. 2).
 FT CONFLICT 462 462 I -> V (IN REF. 3).
 FT CONFLICT 527 527 W -> R (IN REF. 3).
 FT CONFLICT 536 536 S -> M (IN REF. 3).
 SQ SEQUENCE 554 AA; 61123 MW; DFE3F3A3A0520C6B CRC64;

Query Match 69.7%; Score 46; DB 1; Length 554;
 Best Local Similarity 75.0%; Pred. No. 4.4;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 cywgcgyw 8
 ||-||-||
 DB 525 CVWGPQYV 532

RESULT 3
 RELX_HORSE STANDARD; PRT; 182 AA.

AC P22969; Q28907;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Prorelaxin precursor (RXN).
 GN RLN.

OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HOKKAIDO; TISSUE-Placenta;
 RA Min K., Shiota K., Ogawa T.;
 RT "Molecular cloning of equine preprorelaxin cDNA";
 RL J. Reprod. Dev. 42:171-178(1996).
 RN [2]
 RP SEQUENCE OF 32-174 FROM N.A.
 RC TISSUE-Placenta;
 RX MEDLINE=95359320; PubMed=7543295;
 RA Klonisch T., Ryan P.B., Yamashiro S., Porter D.G.;
 RT "Partial complementary deoxyribonucleic acid cloning of equine relaxin messenger ribonucleic acid, and its localization within the equine placenta";
 RL Biol. Reprod. 52:1307-1315(1995).
 RN [3]
 RP SEQUENCE OF 26-53 AND 163-182.

RX MEDLINE=91275796; PubMed=2055195;
 RA Stewart D.R., Nevins B., Hadas E., Vandlen R.;
 RT "Affinity purification and sequence determination of equine relaxin";
 RL Endocrinology 129:375-383(1991).
 CC -!- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN TO PRODUCE DILATATION OF THE BIRTH CANAL IN MANY MAMMALS.
 CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO DISULFIDE BONDS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AB000201; BAA19069.1;
 CC EMBL; S78800; AAB35036.1;
 DR PIR; A49739; A49739.
 DR PIR; B49739; B49739.
 DR HSP; P01348; 4RLX.
 DR InterPro; IPR000739; Insulin_IGF_relaxin.
 DR Pfam; PF00049; Insulin; 1.

DR SMART; SM00078; ILGF; 1.
 DR PROSITE; PS00262; INSULIN; 1.
 KW Insulin family; Hormone; Signal.

FT SIGNAL 1 25
 FT CHAIN 26 53 RELAXIN B CHAIN.
 FT PROPEP 54 156 CONNECTING PEPTIDE.
 FT CHAIN 161 182 RELAXIN A CHAIN.
 FT DISULFID 35 169 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 47 182 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 168 173 BY SIMILARITY.
 FT CONFLICT 66 66 A -> V (IN REF. 2).
 FT CONFLICT 133 133 L -> Q (IN REF. 2).
 SQ SEQUENCE 182 AA; 20721 MW; E5C941430A838B8 CRC64;

Query Match 63.6%; Score 42; DB 1; Length 182;
 Best Local Similarity 100.0%; Pred. No. 6.7;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cywgc 5
 |||||
 DB 169 CYWGC 173

RESULT 4

ID ARSB_RAT STANDARD; PRT; 473 AA.
 AC P50430;

DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Arylsulfatase B (EC 3.1.6.12) (ASB) (N'-acetylgalactosamine-4-sulfatase) (G4S) (Fragment).
 GN ARSB.

OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-Liver;
 RX MEDLINE=96121368; PubMed=8575749;
 RA Kunieda T.;

RT "Mucopolysaccharidosis type VI in rats: isolation of cDNAs encoding arylsulfatase B, chromosomal localization of the gene, and identification of the mutation.";
 RL Genomics 29:582-587(1995).
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 4-sulfate groups of the N-acetyl-D-galactosamine 4-sulfate units of chondroitin sulfate and dermatan sulfate.

CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- DISEASE: DEFECTS IN ARSB ARE THE CAUSE OF MAROTEAUX-LAMY SYNDROME; ALSO KNOWN AS MUCOPOLYSACCHARIDOSIS TYPE VI (MPS-VI).
 CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; D49434; BAA08412.1;
 CC HSP; P15848; 1FSU
 DR InterPro; IPR000917; Sulfatase.
 DR Pfam; PF00884; Sulfatase; 1.
 DR PROSITE; PS00523; SULFATASE_1; 1.
 DR PROSITE; PS00149; SULFATASE_2; FALSE_NEG.
 KW Hydrolase; Glycoprotein; Lysosome; Mucopolysaccharidosis.
 FT NON_TER 1 1
 FT MOD_RES 31 31 2-AMINO-3-OXOPROPIONIC ACID (BY

FT ACT_SITE 87 87 SIMILARITY.
FT DISULFID 57 461 POTENTIAL.
FT BY DISULFID 95 BY SIMILARITY.
FT DISULFID 121 132 BY SIMILARITY.
FT DISULFID 345 387 BY SIMILARITY.
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 366 366 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 473 AA; 53320 MW; 4E114C923A24AF8F CRC64;

Query Match 63.6%; Score 42; DB 1; Length 473;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywgcgyw 8
| | | | |
Db 384 YPGCGYW 390

RESULT 5
ARSB HUMAN
ID ARSB HUMAN STANDARD; PRT; 533 AA.
AC P15848;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Arylsulfatase B precursor (EC 3.1.6.12) (ASB) (N-acetylgalactosamine-4-sulfatase) (G4S).
DE ARSB.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90153994; PubMed=2303452;
RA Peters C., Schmidt B., Komerskirch W., Rupp K., Zuehlendorf M.,
RA Vingron M., Meyer H.E., Pohlmann R., von Figura K.;
RT "Phylogenetic conservation of arylsulfatases. cDNA cloning and
RT expression of human arylsulfatase B.";
RL J. Biol. Chem. 265:3374-3381(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90152677; PubMed=1968043;
RA Schuchman E.H., Jackson C.E., Desnick R.J.;
RT "Human arylsulfatase B: MOPAC cloning, nucleotide sequence of a full-
RT length cDNA, and regions of amino acid identity with arylsulfatases A
RT and C.";
RL Genomics 6:149-158(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93332648; PubMed=7687847;
RA Modaresi S., Rupp K., von Figura K., Peters C.;
RT "Structure of the human arylsulfatase B gene.";
RL Biol. Chem. Hoppe-Seyler 374:327-335(1993).
RN [4]
RP SEQUENCE OF 1-104 FROM N.A.
RX MEDLINE=92028992; PubMed=1930244;
RA Litjens T., Morris C.P., Gibson G.J., Beckmann K.R., Hopwood J.J.;
RT "Human N-acetylgalactosamine-4-sulphatase: protein maturation and
RT isolation of genomic clones.";
RL Biochem. Int. 24:209-215(1991).
RN [5]
RP 2-AMINO-3-OXOPROPIONIC ACID MODIFICATION SITE.
RX MEDLINE=95354208; PubMed=7628016;
RA Schmidt B., Selmer T., Ingendoh A., von Figura K.;
RT "A novel amino acid modification in sulfatases that is defective in
RT multiple sulfatase deficiency.";

RL Cell 82:271-278(1995).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=97184692; PubMed=9032078;
RA Bond C.S., Clements P.R., Ashby S.J., Collyer C.A., Harrop S.J.,
RA Hopwood J.J., Guss J.M.;
RT "Structure of a human lysosomal sulfatase.";
RL Structure 5:277-289(1997).
RN [7]
RP VARIANTS MPS-VI ARG-117; PRO-236 AND TYR-405.
RX MEDLINE=92197625; PubMed=1550123;
RA Jin W.-D., Jackson C.E., Desnick R.J., Schuchman E.H.;
RT "Mucopolysaccharidosis type VI: identification of three mutations in
RT the arylsulfatase B gene of patients with the severe and mild
RT phenotypes provides molecular evidence for genetic heterogeneity.";
RL Am. J. Hum. Genet. 50:795-800(1992).
RN [8]
RP VARIANT MPS-VI VAL-137, AND VARIANT MET-376.
RX MEDLINE=92042029; PubMed=1718978;
RA Wicker G., Prill V., Brooks D., Gibson G., Hopwood J.,
RA von Figura K., Peters C.;
RT "Mucopolysaccharidosis VI (Maroteaux-Lamy syndrome). An intermediate
RT clinical phenotype caused by substitution of valine for glycine at
RT position 137 of arylsulfatase B.";
RL J. Biol. Chem. 266:21386-21391(1991).
RN [9]
RP VARIANTS MPS-VI MET-92; GLN-95; CYS-210; PRO-393 AND PRO-498.
RX MEDLINE=96213747; PubMed=8651289;
RA Litjens T., Brooks D.A., Peters C., Gibson G.J., Hopwood J.J.;
RT "Identification, expression, and biochemical characterization of N-
RT acetylgalactosamine-4-sulfatase mutations and relationship with
RT clinical phenotype in MPS-VI patients.";
RL Am. J. Hum. Genet. 58:1127-1134(1996).
RN [10]
RP VARIANTS MPS-VI TRP-152 AND GLN-160.
RX MEDLINE=94171224; PubMed=8125475;
RA Voskoboieva E., Isbrandt D., von Figura K., Krasnopol'skaya X.,
RA Peters C.;
RT "Four novel mutant alleles of the arylsulfatase B gene in two
RT patients with intermediate form of mucopolysaccharidosis VI
RT (Maroteaux-Lamy syndrome).";
RL Hum. Genet. 93:259-264(1994).
RN [11]
RP VARIANT MPS-IV ARG-302.
RX Villani G.R.D., Balzano N., di Natale P.;
RT "Two novel mutations of the arylsulfatase B gene in two Italian
RT patients with severe form of mucopolysaccharidosis.";
RL Hum. Mutat. 11:410-410(1998).
RN [12]
RP CATALYTIC ACTIVITY: Hydrolysis of the 4-sulfate groups of the N-
RP acetyl-D-galactosamine 4-sulfate units of chondroitin sulfate and
RP dermatan sulfate.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: Lysosomal.
CC -1- DISEASE: DEFECTS IN ARSB ARE THE CAUSE OF MAROTEAUX-LAMY
CC SYNDROME: ALSO KNOWN AS MUCOPOLYSACCHARIDOSIS TYPE VI (MPS-VI).
CC THIS DISEASE IS CHARACTERIZED BY THE ACCUMULATION OF DERMATAN
CC SULFATE IN LYSSOMES. CLINICAL FEATURES CAN INCLUDE ABNORMAL
CC GROWTH, SHORT STATURE, STIFF JOINTS, SKELETAL MALFORMATIONS,
CC CORNEAL CLOUDING, HEPATOSPLENOMEGALY, AND CARDIAC ABNORMALITIES.
CC A WIDE VARIATION IN CLINICAL SEVERITY IS OBSERVED.
CC -1- DISEASE: MULTIPLE SULFATASE DEFICIENCY (MSD) IS A DISORDER THAT
CC COMBINES FEATURES OF METACHROMATIC LEUKODYSTROPHY AND OF
CC MUCOPOLYSACCHARIDOSIS; IT IS CHARACTERIZED BY A DECREASED
CC ACTIVITY OF ALL KNOWN SULFATASES. IT SEEMS TO BE CAUSED FROM THE
CC LACK OF POST-TRANSLATIONAL MODIFICATION OF A CYSTEINE INTO 2-
CC AMINO-3-OXOPROPIONIC ACID.
CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL; J05225; AAA51784.1; -
CC EMBL; M32373; AAA51779.1; -
CC EMBL; X72735; CAA51272.1; -
CC EMBL; X72736; CAA51272.1; JOINED.
CC EMBL; X72737; CAA51272.1; JOINED.
CC EMBL; X72738; CAA51272.1; JOINED.
CC EMBL; X72739; CAA51272.1; JOINED.
CC EMBL; X72740; CAA51272.1; JOINED.
CC EMBL; X72741; CAA51272.1; JOINED.
CC EMBL; X72742; CAA51272.1; JOINED.
CC EMBL; S5777; AAB19988.1; -
CC PIR; A35078; A35078.
CC PDB; 1FSU; 04-FEB-98.
CC MIM; 253200; -
CC MIM; 272200; -
CC InterPro; IPR000917; Sulfatase.
CC Pfam; PF00884; Sulfatase; 1.
CC PROSITE; PS00523; SULFATASE_1; 1.
CC PROSITE; PS00149; SULFATASE_2; 1.
CC Hydrolase; Signal; Glycoprotein; Lysosome; Mucopolysaccharidosis;
KW Disease mutation; Polymorphism; 3D-structure.
FT SIGNAL 1
FT CHAIN 37 533
FT ACT_SITE 147 147
FT MOD_RES 91 91
FT DISULFID 117 521
FT DISULFID 121 155
FT DISULFID 181 192
FT DISULFID 405 447
FT CARBOHYD 188 188
FT CARBOHYD 279 279
FT CARBOHYD 291 291
FT CARBOHYD 366 366
FT CARBOHYD 426 426
FT CARBOHYD 458 458
FT VARIANT 92 92
FT VARIANT 95 95
FT VARIANT 117 117
FT VARIANT 137 137
FT VARIANT 152 152
FT VARIANT 160 160
FT VARIANT 210 210
FT VARIANT 236 236
FT VARIANT 302 302
FT VARIANT 376 376
FT VARIANT 393 393
FT VARIANT 405 405
FT VARIANT 498 498
FT CONFLICT 358 358
FT SEQUENCE 533 AA; 59687 MW; 5983FB6911C4789A CRC64;

Query Match 63.6%; Score 42; DB 1; Length 533;
Best Local Similarity 85.7%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 YWGCYGW 8

```

Db 444 YPCGCGW 450

```

RESULT 6
PHSG_MYCTU STANDARD; PRT; 863 AA.
ID Q10639;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycogen phosphorylase (EC 2.4.1.1).
GN GLGP OR RV1328 OR MT1370 OR MTCY130.13.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch K., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC !- FUNCTION: PHOSPHORYLASE IS AN IMPORTANT ALLOSTERIC ENZYME IN
CC CARBOHYDRATE METABOLISM. ENZYMES FROM DIFFERENT SOURCES DIFFER IN
CC THEIR REGULATORY MECHANISMS AND IN THEIR NATURAL SUBSTRATES.
CC HOWEVER, ALL KNOWN PHOSPHORYLASES SHARE CATALYTIC AND STRUCTURAL
CC PROPERTIES (BY SIMILARITY).
CC !- CATALYTIC ACTIVITY: [(1,4)-alpha-D-glucosyl](N) + phosphate -
CC [(1,4)-alpha-D-glucosyl](N-1) + alpha-D-glucose 1-phosphate.
CC !- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC !- SIMILARITY: BELONGS TO THE GLYCOGEN PHOSPHORYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

EMBL; 273902; CAA98092.1; -
EMBL; AE007010; AA45634.1; -
HSP; P06738; lYGP.
TIGR; MT1370; -
TubercuList; RV1328; -
InterPro; IPR000811; Phosphorylase.
PROSITE; PS00102; PHOSPHORYLASE; 1.
Transferase; Glycosyltransferase; Carbohydrate metabolism;
KW Glycogen metabolism; Pyridoxal phosphate; Complete proteome.
FT BINDING 618 618 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT CONFLICT 255 255 K -> E (IN REF. 2).
FT SEQUENCE 863 AA; 95515 MW; E39826B03D5B374F CRC64;

```

Query Match 63.6%; Score 42; DB 1; Length 863;
Best Local Similarity 71.4%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ywgcgyw 8
|||||

Db 595 YWGCDDW 601

RESULT 7

ACM3_MOUSE
ID ACM3_MOUSE STANDARD; PRT; 589 AA.

AC Q9ERZ3; Q64055;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Muscarinic acetylcholine receptor M3 (Mm3 mAChR).

GN CHRM3 OR CHRM-3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Gomez J., Wess J.;

RT "Isolation, sequence and functional expression of mouse muscarinic

RT acetylcholine receptor genes.";

RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 314-439 FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=95179320; PubMed=7874308;

RA Andre C., Dos Santos G., Koulakoff A.;

RT "Cultured neurons from mouse brain reproduce the muscarinic receptor

RT profile of their tissue of origin.";

RL Eur. J. Neurosci. 6:1691-1701(1994).

CC -!- FUNCTION: THE MUSCARINIC ACETYLCHOLINE RECEPTOR MEDIATES VARIOUS

CC CELLULAR RESPONSES, INCLUDING INHIBITION OF ADENYLATE CYCLASE,

CC BREAKDOWN OF PHOSPHOINOSITIDES & MODULATION OF POTASSIUM CHANNELS

CC THROUGH THE ACTION OF G PROTEINS. PRIMARY TRANSDUCING EFFECT IS PI

CC TURNOVER.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF264050; AAG14344.1; ..

DR EMBL; S74908; AAB33576.2; ..

DR MGD; MGI:88398; Chrm3.

DR InterPro; IPR000276; GPCR_Rhodpsn.

DR Pfam; PF00001; 7tm_1; 1.

DR PRINTS; PR00237; GPCRRHODPSN.

DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; 1.

DR PROSITE; PS50262; G-PROTEIN_RECEP_FL_2; 1.

KW Postsynaptic membrane; Ionic channel; Glycoprotein; Transmembrane;

KW Phosphorylation; Multigene family; G-protein coupled receptor.

FT DOMAIN 1 66 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 67 90 1 (POTENTIAL).

FT DOMAIN 91 103 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 104 124 2 (POTENTIAL).

FT DOMAIN 125 141 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 142 163 3 (POTENTIAL).

FT DOMAIN 164 183 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 184 206 4 (POTENTIAL).

FT DOMAIN 207 228 EXTRACELLULAR (POTENTIAL).

FT

FT TRANSMEM 229 251 5 (POTENTIAL).

FT DOMAIN 252 491 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 492 512 6 (POTENTIAL).

FT DOMAIN 513 526 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 527 546 7 (POTENTIAL).

FT DOMAIN 547 589 CYTOPLASMIC (POTENTIAL).

FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT DISULFID 140 220 BY SIMILARITY.

SQ SEQUENCE 589 AA; 66211 MW; AB782149EBEE7804 CRC64;

Query Match 62.1%; Score 41; DB 1; Length 589;

Best Local Similarity 71.4%; Pred. No. 26;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ywgcgyw 8

|||||

Db 524 YWNLGYW 530

RESULT 8

ACM3_RAT

ID ACM3_RAT STANDARD; PRT; 589 AA.

AC P08483; Q9QWK9;

DT 01-AUG-1988 (Rel. 08, Created)

DT 01-NOV-1988 (Rel. 09, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Muscarinic acetylcholine receptor M3.

GN CHRM3 OR CHRM-3.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=87263421; PubMed=3037705;

RA Bonner T.I., Buckley N.J., Young A.C., Brann M.R.;

RT "Identification of a family of muscarinic acetylcholine receptor

RT genes.";

RL Science 237:527-532(1987).

RN [2]

RP REVISION TO 184.

RX MEDLINE=90166521; PubMed=3272174;

RA Bonner T.I., Young A.C., Brann M.R., Buckley N.J.;

RT "Cloning and expression of the human and rat m5 muscarinic

RT acetylcholine receptor genes.";

RL Neuron 1:403-410(1988).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=88077068; PubMed=3120722;

RA Braun T., Schofield P.R., Shivers B.D., Pritchett D.B., Seeburg P.H.;

RT "A novel subtype of muscarinic receptor identified by homology

RT screening.";

RL Biochem. Biophys. Res. Commun. 149:125-132(1987).

RN [4]

RP SEQUENCE FROM N.A.

RC TISSUE=Iris;

RX MEDLINE=99138467; PubMed=9972520;

RA Furuta M., Ohya S., Imaizumi Y., Watanabe M.;

RT "Molecular cloning of m3 muscarinic acetylcholine receptor in rat

RT iris.";

RL J. Smooth Muscle Res. 34:111-122(1998).

RN [5]

RP MUTAGENESIS.

RX MEDLINE=92037535; PubMed=1657592;

RA Wess J., Gdula D., Brann M.R.;

RT "Site-directed mutagenesis of the m3 muscarinic receptor:

RT identification of a series of threonine and tyrosine residues

```
RT involved in agonist but not antagonist binding.";
RL EMBO J. 10:3729-3734(1991).
RN [6]
RP MUTAGENESIS.
RX MEDLINE-92406875; PubMed-1527051;
RA Wess J., Magglio R., Palmer J.R., Vogel Z.;
RT "Role of conserved threonine and tyrosine residues in acetylcholine
RT binding and muscarinic receptor activation. A study with m3
RT muscarinic receptor point mutants.";
RL J. Biol. Chem. 267:19313-19319(1992).
CC -!- FUNCTION: THE MUSCARINIC ACETYLCHOLINE RECEPTOR MEDIATES VARIOUS
CC CELLULAR RESPONSES, INCLUDING INHIBITION OF ADENYLATE CYCLASE,
CC BREAKDOWN OF PHOSPHOINOSITIDES & MODULATION OF POTASSIUM CHANNELS
CC THROUGH THE ACTION OF G PROTEINS. PRIMARY TRANSDUCING EFFECT IS PI
CC TURNOVER.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M16407; AAA40661.1; ALT SEQ.
CC EMBL; M16408; AAA40662.1; ALT SEQ.
CC EMBL; M18088; AAA40659.1; -.
CC EMBL; M62826; AAA41553.1; -.
CC EMBL; AB017656; BAA36839.1; -.
CC PIR; B29514; B29514.
CC GCRdb; GCR_0116; -.
CC GCRdb; GCR_0118; -.
CC GCRdb; GCR_0119; -.
CC GCRdb; GCR_0140; -.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHHODPSN.
CC PROSITE; PS00237; G_PROTEIN_RECEPT_F1_1; 1.
CC PROSITE; PS00262; G_PROTEIN_RECEPT_F1_2; 1.
CC Postsynaptic membrane; Ionic channel; Glycoprotein; Transmembrane;
CC Phosphorylation; Multigene family; G-protein coupled receptor.
CC -----
CC DOMAIN 1 66 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 67 90 1 (POTENTIAL).
CC DOMAIN 91 103 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 104 124 2 (POTENTIAL).
CC DOMAIN 125 141 3 (POTENTIAL).
CC TRANSMEM 142 163 EXTRACELLULAR (POTENTIAL).
CC DOMAIN 164 183 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 184 206 EXTRACELLULAR (POTENTIAL).
CC DOMAIN 207 228 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 229 251 5 (POTENTIAL).
CC DOMAIN 252 491 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 492 512 6 (POTENTIAL).
CC DOMAIN 513 526 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 527 546 7 (POTENTIAL).
CC DOMAIN 547 589 CYTOPLASMIC (POTENTIAL).
CC CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC DISULFID 140 220 BY SIMILARITY.
CC CONFLICT 184 184 A -> R (IN REF. 4).
CC CONFLICT 516 516 C -> R (IN REF. 3).
CC CONFLICT 556 556 T -> M (IN REF. 3).
CC SEQUENCE 589 AA; 66065 MW; 9A5EF2FA653C830A CRC64;
Query Match 62.1%; Score 41; DB 1; Length 589;
Best Local Similarity 71.4%; Pred. No. 26;
```

```
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 YWGCYGYW 8
DB 524 YWNLYGYW 530
RESULT 9
ACM3_BOVIN
ID ACM3_BOVIN STANDARD; PRT; 590 AA.
AC P41984;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Muscarinic acetylcholine receptor M3.
GN CHRM3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE-94339178; PubMed-8061048;
RA Lee P.H., Hodges P.K., Glickman F., Chang K.J.;
RT "Cloning and expression of a cDNA encoding bovine muscarinic
RT acetylcholine m3 receptor.";
RL Biochim. Biophys. Acta 1223:151-154(1994).
CC [2]
CC SEQUENCE OF 327-467 FROM N.A.
CC TISSUE=Adrenal gland;
CC Sui A.-L., Chou W.-Y., Kao L.-S.;
CC Submitted (XXX-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THE MUSCARINIC ACETYLCHOLINE RECEPTOR MEDIATES VARIOUS
CC CELLULAR RESPONSES, INCLUDING INHIBITION OF ADENYLATE CYCLASE,
CC BREAKDOWN OF PHOSPHOINOSITIDES & MODULATION OF POTASSIUM CHANNELS
CC THROUGH THE ACTION OF G PROTEINS. PRIMARY TRANSDUCING EFFECT IS PI
CC TURNOVER.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U08286; AAA51866.1; -.
CC EMBL; L27103; AAA30653.1; -.
CC GCRdb; GCR_0849; -.
CC GCRdb; GCR_1227; -.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHHODPSN.
CC PROSITE; PS00237; G_PROTEIN_RECEPT_F1_1; 1.
CC PROSITE; PS00262; G_PROTEIN_RECEPT_F1_2; 1.
CC Postsynaptic membrane; Ionic channel; Glycoprotein; Transmembrane;
CC Phosphorylation; Multigene family; G-protein coupled receptor.
CC -----
CC DOMAIN 1 67 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 68 91 1 (POTENTIAL).
CC DOMAIN 92 104 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 105 125 2 (POTENTIAL).
CC DOMAIN 126 142 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 143 164 3 (POTENTIAL).
CC DOMAIN 165 184 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 185 207 4 (POTENTIAL).
CC DOMAIN 208 229 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 230 252 5 (POTENTIAL).
CC DOMAIN 253 492 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 493 513 6 (POTENTIAL).
```


FT DOMAIN 514 527 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 528 547 7 (POTENTIAL).
FT DOMAIN 548 590 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 7 7 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 141 221 BY SIMILARITY.
FT CONFLICT 424 424 F -> S (IN REF. 2).
FT CONFLICT 438 438 A -> G (IN REF. 2).
FT CONFLICT 440 440 A -> G (IN REF. 2).
FT CONFLICT 452 452 A -> G (IN REF. 2).
FT CONFLICT 461 461 A -> G (IN REF. 2).
FT CONFLICT 467 467 F -> L (IN REF. 2).
SQ SEQUENCE 590 AA; 66103 MW; 4DE04EDE33CCA8D6 CRC64;

Query Match 62.1%; Score 41; DB 1; Length 590;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ywgcgyw 8

Db 525 YWNLGYW 531

RESULT 10
ACM3_PIG STANDARD; PRT; 590 AA.
ID ACM3_PIG
AC P11483;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Muscarinic acetylcholine receptor M3.
GN CHR3.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88296835; PubMed=3402600;
RA Akiba I., Kubo T., Maeda A., Bujo H., Nakai J., Mishina M., Numa S.;
RT "Primary structure of porcine muscarinic acetylcholine receptor III
and antagonist binding studies.";
RL FEBS Lett. 235:257-261(1988).
CC -!- FUNCTION: THE MUSCARINIC ACETYLCHOLINE RECEPTOR MEDIATES VARIOUS
CELLULAR RESPONSES, INCLUDING INHIBITION OF ADENYLATE CYCLASE,
BREAKDOWN OF PHOSPHOINOSITIDES & MODULATION OF POTASSIUM CHANNELS
THROUGH THE ACTION OF G PROTEINS. PRIMARY TRANSDUCING EFFECT IS PI
TURNOVER.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X12712; CAA31215.1; -.
CC GCD: S01114; S01114.
CC GCRD: GCR_0104; -.
CC InterPro: IPR000276; GPCR_Rhodopsn.
CC Pfam: PF00001; 7tm1; 1.
CC PRINTS: P000237; GPCR_Rhodopsn.
CC PROSITE: P500237; G-PROTEIN_RECEPTOR_F1_1; 1.
CC PROSITE: P500262; G-PROTEIN_RECEPTOR_F1_2; 1.
CC Postsynaptic membrane; Ionic channel; Glycoprotein; Transmembrane;

KW Phosphorylation; Multigene family; G-protein coupled receptor.
FT DOMAIN 1 67 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 68 91 1 (POTENTIAL).
FT DOMAIN 92 104 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 105 125 2 (POTENTIAL).
FT DOMAIN 126 142 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 143 164 3 (POTENTIAL).
FT DOMAIN 165 184 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 185 207 4 (POTENTIAL).
FT DOMAIN 208 229 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 230 252 5 (POTENTIAL).
FT DOMAIN 253 492 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 493 513 6 (POTENTIAL).
FT DOMAIN 514 527 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 528 547 7 (POTENTIAL).
FT DOMAIN 548 590 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 7 7 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 8 8 O-LINKED (POTENTIAL).
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 141 221 BY SIMILARITY.
SQ SEQUENCE 590 AA; 66077 MW; 9998D2A4802FD32A CRC64;

Query Match 62.1%; Score 41; DB 1; Length 590;

Best Local Similarity 71.4%; Pred. No. 26;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ywgcgyw 8

Db 525 YWNLGYW 531

RESULT 11
YD35_MYCPN STANDARD; PRT; 741 AA.
ID YD35_MYCPN
AC P75443;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein MPN335 (F10_orf741).
GN MPN335 OR MP501.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfeich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Hermann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
pneumoniae";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO M.PNEUMONIAE MPN333.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AE000049; AAB96149.1; -.
CC InterPro: IPR000188; GABAA_receptor.
CC Pfam: PF02932; Neur_chan_membr; 1.
CC Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 34 54 POTENTIAL.
 FT TRANSMEM 76 96 POTENTIAL.
 FT TRANSMEM 120 140 POTENTIAL.
 FT TRANSMEM 156 176 POTENTIAL.
 FT TRANSMEM 187 207 POTENTIAL.
 FT TRANSMEM 655 675 POTENTIAL.
 FT TRANSMEM 715 735 POTENTIAL.
 SQ SEQUENCE 741 AA; 85025 MW; 1776A96BCF83567F CRC64;

Query Match 62.1%; Score 41; DB 1; Length 741;
 Best Local Similarity 71.4%; Pred. No. 31;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 ywgcyyw 8
 |||||
 Db 280 YWNVGYW 286

RESULT 12
 SAP_HUMAN
 ID SAP_HUMAN STANDARD; PRT; 524 AA.
 AC P07602; P07292; P15793; P78538; P78546; P78547; Q92741; Q92742;
 AC Q92740; Q92739; P78541; P78558;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Proactivator polypeptide precursor [Contains: Saposin A (Protein A);
 DE Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside
 DE sulfate activator) (CSAct) (Dispersin) (Sulfatide/GM1 activator);
 DE Saposin C (Co-beta-glucosidase) (A1 activator) (Glucosylceramidase
 DE activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D
 DE (Protein C) (Component C)].
 GN PSAP.
 GN PSAP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE-90129043; PubMed-25151510;
 RA Rorman E.G., Grabowski G.A.;
 RT "Molecular cloning of a human co-beta-glucosidase cDNA: evidence that
 RT four sphingolipid hydrolase activator proteins are encoded by single
 RT genes in humans and rats.";
 RL Genomics 5:486-492(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89255151; PubMed-2498298;
 RA Nakano T., Sandhoff K., Stuenkel J., Christomanou H., Suzuki K.;
 RT "Structure of full-length cDNA coding for sulfatide activator, a
 RT Co-beta-glucosidase and two other homologous proteins: two alternate
 RT forms of the sulfatide activator.";
 RL J. Biochem. 105:152-154(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Skin, and Eye;
 RC Strausberg R.;
 RA Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 59-125 AND 304-513 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE-91192146; PubMed-2013321;
 RA Holtschmidt H., Sandhoff K., Fuerst W., Kwon H.Y., Schnabel D.,
 RA Suzuki K.;
 RT "The organization of the gene for the human cerebroside sulfate
 RT activator protein.";
 RL FEBS Lett. 280:267-270(1991).
 RN [5]
 RP SEQUENCE OF 164-524 FROM N.A.
 RX MEDLINE-88068647; PubMed-2825202;
 RA Dewji N.N., Wenger D.A., O'Brien J.S.;

RT "Nucleotide sequence of cloned cDNA for human sphingolipid activator
 RT protein 1 precursor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8652-8656(1987).
 RN [6]
 RP PARTIAL SEQUENCE OF 60-142.
 RX MEDLINE-89240739; PubMed-2717620;
 RA Morimoto S., Martin B.M., Yamamoto Y., Kretz K.A., O'Brien J.S.,
 RA Kishimoto Y.;
 RT "Saposin A: second cerebroside activator protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3389-3393(1989).
 RN [7]
 RP SEQUENCE OF 195-263 FROM N.A.
 RX MEDLINE-86130593; PubMed-2668718;
 RA Dewji N.N., Wenger D.A., Fujibayashi S., Donoviel M., Esch F.,
 RA Hill F., O'Brien J.S.;
 RT "Molecular cloning of the sphingolipid activator protein-1 (SAP-1),
 RT the sulfatide sulfatase activator.";
 RL Biochem. Biophys. Res. Commun. 134:989-994(1986).
 RN [8]
 RP SEQUENCE OF 195-274.
 RC TISSUE=Kidney;
 RX MEDLINE-91006165; PubMed-2209618;
 RA Furst W., Schubert J., Machleidt W., Meyer H.E., Sandhoff K.;
 RT "The complete amino-acid sequences of human ganglioside GM2 activator
 RT protein and cerebroside sulfate activator protein.";
 RL Eur. J. Biochem. 192:709-714(1990).
 RN [9]
 RP SEQUENCE OF 195-274.
 RX MEDLINE-89207118; PubMed-3242555;
 RA Kleinschmidt T., Christomanou H., Braunitzer G.;
 RT "Complete amino-acid sequence of the naturally occurring A2 activator
 RT protein for enzymic sphingomyelin degradation: identity to the
 RT sulfatide activator protein (SAP-1).";
 RL Biol. Chem. Hoppe-Seyler 369:1361-1365(1988).
 RN [10]
 RP SEQUENCE OF 311-390.
 RX MEDLINE-88163077; PubMed-3442600;
 RA Kleinschmidt T., Christomanou H., Braunitzer G.;
 RT "Complete amino-acid sequence and carbohydrate content of the
 RT naturally occurring glucosylceramide activator protein (A1 activator)
 RT absent from a new human Gaucher disease variant.";
 RL Biol. Chem. Hoppe-Seyler 368:1571-1578(1987).
 RN [11]
 RP SEQUENCE OF 407-484.
 RX MEDLINE-89000190; PubMed-3048308;
 RA Furst W., Machleidt W., Sandhoff K.;
 RT "The precursor of sulfatide activator protein is processed to three
 RT different proteins.";
 RL Biol. Chem. Hoppe-Seyler 369:317-328(1988).
 RN [12]
 RP PARTIAL SEQUENCE OF 405-484.
 RX MEDLINE-89025876; PubMed-2845979;
 RA Morimoto S., Martin B.M., Kishimoto Y., O'Brien J.S.;
 RT "Saposin D: a sphingomyelinase activator.";
 RL Biochem. Biophys. Res. Commun. 156:403-410(1988).
 RN [13]
 RP SEQUENCE OF 17-26.
 RC TISSUE=Milk;
 RX MEDLINE-92068206; PubMed-1958198;
 RA Kondoh K., Hineno T., Sano A., Kakimoto Y.;
 RT "Isolation and characterization of prosaposin from human milk.";
 RL Biochem. Biophys. Res. Commun. 181:286-292(1991).
 RN [14]
 RP PARTIAL SEQUENCE OF SAPOSIN B, AND STRUCTURE OF CARBOHYDRATES.
 RC TISSUE=Urine;
 RX MEDLINE-20032116; PubMed-10562467;
 RA Fluharty A.L., Lombardo C., Louis A., Stevens R.L., Whitelegge J.,
 RA Waring A.J., To T., Fluharty C.B., Faull K.F.;
 RT "Preparation of the cerebroside sulfate activator (CSAct or saposin B)
 RT from human urine.";
 RL Mol. Genet. Metab. 68:391-403(1999).
 RN [15]
 RP STRUCTURE OF CARBOHYDRATE ON ASN-215.

RA MEDLINE=21110404; PubMed=11180632;
RA Fauli K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,
RA Stevens R.L., Fluharty C.B., Fluharty A.L.;
RT "Structure of the asparagine-linked sugar chains of porcine kidney and
RT human urine cerebroside sulfate activator protein.";
RL J. Mass Spectrom. 35:1416-1424(2000).
RN [16]
RN MASS SPECTROMETRY.
RC TISSUE=Urine;
RX MEDLINE=99411404; PubMed=10510427;
RA Fauli K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,
RA Krutchinsky A.N., Standing K.G., Waring A.J., Stevens R.L.,
RA Fluharty C.B., Fluharty A.L.;
RT "Cerebroside sulfate activator protein (Saposin B): chromatographic
RT and electrospray mass spectrometric properties.";
RL J. Mass Spectrom. 34:1040-1054(1999).
RN [17]
RN VARIANT MLD ILE-217.
RX MEDLINE=90147748; PubMed=2302219;
RA Rafi M.A., Zhang X.-L., Degala G., Wenger D.A.;
RT "Detection of a point mutation in sphingolipid activator protein-1
RT mRNA in patients with a variant form of metachromatic
RT leukodystrophy.";
RL Biochem. Biophys. Res. Commun. 166:1017-1023(1990).
RN [18]
RN SEQUENCE FROM N.A., AND VARIANT MLD ILE-217.
RX MEDLINE=90207231; PubMed=2320574;
RA Kretz K.A., Carson G.S., Morimoto S., Kishimoto Y., Fluharty A.L.,
RA O'Brien J.S.;
RT "Characterization of a mutation in a family with saposin B
RT deficiency: a glycosylation site defect.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2541-2544(1990).
RN [19]
RN VARIANT MLD SER-241, SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=91210267; PubMed=2019586;
RA Holtschmidt H., Sandhoff K., Kwon H.Y., Harzer K., Nakano T.,
RA Suzuki K.;
RT "Sulfatide activator protein. Alternative splicing that generates
RT three mRNAs and a newly found mutation responsible for a clinical
RT disease.";
RL J. Biol. Chem. 266:7556-7560(1991).
RN [20]
RN VARIANT GAUCHER PHE-388.
RX MEDLINE=91285107; PubMed=2060627;
RA Schnabel D., Schroeder M., Sandhoff K.;
RT "Mutation in the sphingolipid activator protein 2 in a patient with a
RT variant of Gaucher disease.";
RL FEBS Lett. 284:57-59(1991).
RN [21]
RN REVIEW ON MLD VARIANTS.
RX MEDLINE=95107731; PubMed=7866401;
RA Gieselmann V., Zlotogora J., Harris A., Wenger D.A., Morris C.P.;
RT "Molecular genetics of metachromatic leukodystrophy.";
RL Hum. Mutat. 4:233-242(1994).
CC -!- FUNCTION: The lysosomal degradation of sphingolipids takes place
CC by the sequential action of specific hydrolases. Some of these
CC enzymes require specific low-molecular mass, non-enzymic proteins:
CC the sphingolipids activator proteins (coproteins).
CC -!- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
CC galactosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
CC Saposin C apparently acts by combining with the enzyme and acidic
CC lipid to form an activated complex, rather than by solubilizing
CC the substrate.
CC -!- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1
CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
CC Saposin B forms a solubilizing complex with the substrates of the
CC sphingolipid hydrolases.
CC -!- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase
CC activator (EC 3.1.4.12).
CC -!- SUBUNIT: Saposin B is a homodimer.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC -----

CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- ALTERNATIVE PRODUCTS: At least 3 isoforms; sap-mu-0 (shown here),
Query Match 60.6%; Score 40; DB 1; Length 524;
Best Local Similarity 62.5%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 cywgcgyw 8
Db 495 CIWGPSYW 502
RESULT 13
ID SAP_BOVIN STANDARD; PRT; 525 AA.
AC P26779; Q9NZG4;
DT 01-AUG-1992 (Rel. 23, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Proactivator polypeptide precursor [Contains: Saposin A (Protein A);
DE Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside
DE sulfate activator) (CSAct) (Dispersin) (Sulfatide/GM1 activator);
DE Saposin C (Co-beta-glucosidase) (A1 activator) (Glucosylceramidase
DE activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D
DE (Protein C) (Component C)].
GN PSAP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS.
RC TISSUE=Mammary gland;
RA Azuma N., Yoshida K.;
RT "RT-PCR cloning of bovine prosaposin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 312-391.
RC TISSUE=Spleen;
RX MEDLINE=92207994; PubMed=1554743;
RA Sano A., Mizuno T., Kondoh K., Hinenio T., Ueno S.-I., Kakimoto Y.,
RA Morita N.;
RT "Saposin-C from bovine spleen; complete amino acid sequence and
RT relation between the structure and its biological activity.";
RL Biochim. Biophys. Acta 1120:75-80(1992).
CC -!- FUNCTION: THE LYSSOMAL DEGRADATION OF SPHINGOLIPIDS TAKES PLACE
CC BY THE SEQUENTIAL ACTION OF SPECIFIC HYDROLASES. SOME OF THESE
CC ENZYMES REQUIRE SPECIFIC LOW-MOLECULAR MASS, NON-ENZYMIC PROTEINS:
CC THE SPHINGOLIPIDS ACTIVATOR PROTEINS (COPROTEINS) (BY SIMILARITY).
CC -!- FUNCTION: SAPOSIN A AND SAPOSIN C STIMULATE THE HYDROLYSIS OF
CC GLUCOSYLCERAMIDE BY BETA-GLUCOSYLCERAMIDASE (EC 3.2.1.45) AND
CC GALACTOSYLCERAMIDE BY BETA-GALACTOSYLCERAMIDASE (EC 3.2.1.46).
CC SAPOSIN-C APPARENTLY ACTS BY COMBINING WITH THE ENZYME AND ACIDIC
CC LIPID TO FORM AN ACTIVATED COMPLEX, RATHER THAN BY SOLUBILIZING
CC THE SUBSTRATE.
CC -!- FUNCTION: SAPOSIN B STIMULATES THE HYDROLYSIS OF GALACTO-
CC CEREBROSIDE SULFATE BY ARYLSULFATASE A (EC 3.1.6.8), GM1
CC GANGLIOSIDES BY BETA-GALACTOSIDASE (EC 3.2.1.23) AND
CC GLOBOTRIAOSYLCERAMIDE BY ALPHA-GALACTOSIDASE A (EC 3.2.1.22).
CC SAPOSIN B FORMS A SOLUBILIZING COMPLEX WITH THE SUBSTRATES OF THE
CC SPHINGOLIPID HYDROLASES (BY SIMILARITY).
CC -!- FUNCTION: SAPOSIN D IS A SPECIFIC SPHINGOMYELIN PHOSPHODIESTERASE
CC ACTIVATOR (EC 3.1.4.12) (BY SIMILARITY).
CC -!- SUBUNIT: SAPOSIN B IS A HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- PTM: THIS PRECURSOR IS PROTEOLYTICALLY PROCESSED TO 4 SMALL
CC PEPTIDES, WHICH ARE SIMILAR TO EACH OTHER AND ARE SPHINGOLIPID
CC HYDROLASE ACTIVATOR PROTEINS (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 2 SAPOSIN A-TYPE DOMAINS.
CC -!- SIMILARITY: CONTAINS 4 SAPOSIN B-TYPE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC -----

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AB036791; BAA95677.1; -;
PIR; S21770.
InterPro; IPR003119; Sapa.
InterPro; IPR000004; SapB.
InterPro; IPR003259; Saposin.
Pfam; PF02199; SAPA; 2.
ProDom; PD001732; SapB; 3.
ProDom; PD012321; Saposin; 1.
SMART; SM001162; Saposin; 2.
SMART; SM00118; SAPA; 4.
Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat.

FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 60 142 SAPOSIN A.
FT CHAIN 196 275 SAPOSIN B.
FT CHAIN 312 392 SAPOSIN C.
FT CHAIN 406 487 SAPOSIN D.
FT DOMAIN 21 54 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 59 142 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 194 276 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 312 393 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 406 487 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 492 525 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 63 138 BY SIMILARITY.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 199 272 BY SIMILARITY.
FT DISULFID 202 266 BY SIMILARITY.
FT DISULFID 231 242 BY SIMILARITY.
FT DISULFID 316 389 BY SIMILARITY.
FT DISULFID 319 383 BY SIMILARITY.
FT DISULFID 347 358 BY SIMILARITY.
FT DISULFID 410 483 BY SIMILARITY.
FT DISULFID 413 477 BY SIMILARITY.
FT DISULFID 441 452 BY SIMILARITY.
FT CARBOHYD 80 80 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 101 101 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 427 427 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 127 127 H -> R.
FT VARIANT 260 263 MQPK -> IRIR.
FT CONFLICT 317 317 E -> Q (IN REF. 2).
FT CONFLICT 367 367 R -> S (IN REF. 1).
SEQUENCE 525 AA; 58120 MW; 293AF0FB9C4FA99 CRC64;

Query Match 60.6%; Score 40; DB 1; Length 525;
Best Local Similarity 62.5%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 cywgcgyw 8
| | | |
Db 496 CVWGPSTW 503

RESULT 14
SAP_MOUSE
ID SAP_MOUSE STANDARD; PRT; 557 AA.
AC O61207; O64219; O64006; O60861;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DR Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
GN PSAP OR SGP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;
RA Tsuda M., Sakiyama T., Endo H., Kitagawa T.;
RL "The primary structure of mouse saposin.";
Biochem. Biophys. Res. Commun. 184:1266-1272(1992).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94272718; PubMed=1590788;
RA Tsuda M., Sakiyama T., Endo H., Kitagawa T.;
RL "The primary structure of mouse saposin.";
Biochem. Biophys. Res. Commun. 184:1266-1272(1992).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94272317; PubMed=8003952;
RA Sprecher-Levy H., Orr-Urtreger A., Lonai P., Horowitz M.;
RL "Murine prosaposin: expression in the reproductive system of a gene implicated in human genetic disease.";
Cell. Mol. Biol. 40:233-233(1994).
[4]
RP SEQUENCE FROM N.A.
RX MEDLINE=96084310; PubMed=8565332;
RA Cao Q.P., Crain W.R.;
RL "Expression of SGP-1 mRNA in preimplantation mouse embryos.";
Dev. Genet. 17:263-271(1995).
[5]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C; TISSUE=Liver;
RA Zhao Q.O., Hay N.N., Morales C.R.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: CONTAINS 2 SAPOSIN A-TYPE DOMAINS.
CC -!- SIMILARITY: CONTAINS 4 SAPOSIN B-TYPE DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; S36200; AAB22175.1; -;
EMBL; S71616; AAB31059.1; -;
EMBL; U27340; AAB92567.1; -;
EMBL; U57999; AAB02695.1; -;
MGD; MGI:97783; Psap.
InterPro; IPR003119; Sapa.
InterPro; IPR000004; SapB.
InterPro; IPR003259; Saposin.
Pfam; PF02199; SAPA; 2.
ProDom; PD001732; SapB; 3.
ProDom; PD012321; Saposin; 1.
SMART; SM00162; SAPA; 2.
SMART; SM00118; SAPA; 4.
Sulfation; Signal; Glycoprotein; Repeat.

KW SIGNAL 1 16 BY SIMILARITY.
FT CHAIN 17 557 SULFATED GLYCOPROTEIN 1.
FT DOMAIN 21 54 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 59 142 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 193 277 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 313 394 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 438 519 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 524 557 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 63 138 BY SIMILARITY.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 197 273 BY SIMILARITY.
FT DISULFID 200 267 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 317 390 BY SIMILARITY.
FT DISULFID 320 384 BY SIMILARITY.
FT DISULFID 348 359 BY SIMILARITY.
FT DISULFID 442 515 BY SIMILARITY.
FT DISULFID 445 509 BY SIMILARITY.
FT DISULFID 473 484 BY SIMILARITY.

```
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 334 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 334 334 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 83 83 Q -> E (IN REF. 2).
FT CONFLICT 158 158 I -> V (IN REF. 3).
FT CONFLICT 160 160 MISSING (IN REF. 2).
FT CONFLICT 171 172 MS -> SA (IN REF. 3).
FT CONFLICT 244 244 V -> L (IN REF. 2).
FT CONFLICT 254 254 M -> I (IN REF. 3).
FT CONFLICT 255 255 L -> W (IN REF. 2).
FT CONFLICT 260 262 MISSING (IN REF. 3).
FT CONFLICT 307 307 N -> D (IN REF. 2).
FT CONFLICT 322 322 F -> L (IN REF. 2).
FT CONFLICT 349 350 AL -> GV (IN REF. 1).
FT CONFLICT 367 367 G -> D (IN REF. 3).
FT CONFLICT 370 370 L -> Q (IN REF. 2).
FT CONFLICT 373 373 I -> D (IN REF. 3).
FT CONFLICT 391 391 A -> T (IN REF. 3).
FT CONFLICT 393 393 R -> L (IN REF. 3).
FT CONFLICT 406 406 A -> R (IN REF. 2 AND 3).
FT CONFLICT 430 430 P -> R (IN REF. 2).
FT CONFLICT 445 445 C -> F (IN REF. 3).
FT CONFLICT 448 448 L -> P (IN REF. 4).
SQ SEQUENCE 557 AA; 61422 MW; 134593E20499E35E CRC64;

Query Match 60.6%; Score 40; DB 1; Length 557;
Best Local Similarity 62.5%; Pred. No. 35;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 cywgcgyw 8
   | | | | |
Db 528 CVWGPSYW 535

RESULT 15
ID INAL TRIHA STANDARD; PRT; 573 AA.
AC P34054;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Amino-acid permease INDAL.
GN INDAL.
OS Trichoderma harzianum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; mitosporic Hypocreales; Trichoderma.
OX NCBI_TaxID=5544;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IMI 206040;
RX MEDLINE=95291429; PubMed=7773384;
RA Vasseur V.V., van Montagu M.M., Goldman G.G.H.:
RT "Trichoderma harzianum genes induced during growth on Rhizoctonia
   solani cell walls.";
RL Microbiology 141:767-774(1995).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- INDUCTION: DURING MYCOPARASITISM.
CC -!- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; 222594; CAAR0308.1; -.
DR PIR; S33212; S33212.
DR InterPro; IPR002293; AA_rel_permease.1.
DR InterPro; IPR002027; Amino_acid_permease.
```

```
DR Pfam; PF00324; aa_permeases; 1.
DR PROSITE; PS00218; AMINO_ACID_PERMEASE_1; 1.
KW Transport; Amino-acid transport; Transmembrane.
FT TRANSMEM 72 92 POTENTIAL.
FT TRANSMEM 99 117 POTENTIAL.
FT TRANSMEM 176 200 POTENTIAL.
FT TRANSMEM 212 229 POTENTIAL.
FT TRANSMEM 257 280 POTENTIAL.
FT TRANSMEM 296 315 POTENTIAL.
FT TRANSMEM 351 371 POTENTIAL.
FT TRANSMEM 398 415 POTENTIAL.
FT TRANSMEM 425 444 POTENTIAL.
FT TRANSMEM 470 494 POTENTIAL.
FT TRANSMEM 511 527 POTENTIAL.
SQ SEQUENCE 573 AA; 62850 MW; 5F80A806934DB55D CRC64;

Query Match 60.6%; Score 40; DB 1; Length 573;
Best Local Similarity 71.4%; Pred. No. 35;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 YWgcgyw 8
   | | | | |
Db 239 YWGARYW 245

Search completed: August 12, 2002, 17:27:26
Job time: 338 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:28:14 ; Search time 41.64 Seconds
(without alignments)
33.236 Million cell updates/sec

Title: 10-071247-2

Perfect score:

Sequence: 1 cywgcgYW 8

Scoring table: BLOSUM62

scoring table. BBO30M02
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Post-processing: Minimum Match 100%

Maximum Match 100%
Listing first 45 summaries

Database :

```

1: 1: sp_archaea:*
2: 1: sp_archaea:*
3: 2: sp_bacteria:*
4: 3: sp_fungi:*
5: 4: sp_human:*
6: 5: sp_invertebrate:*
7: 6: sp_mammal:*
8: 7: sp_mhc:*
9: 8: sp_organelle:*
10: 9: sp_phage:*
11: 10: sp_plant:*
12: 11: sp_rodent:*
13: 12: sp_virus:*
14: 13: sp_vertebrate:*
15: 14: sp_unclassified:*
16: 15: sp_rvirus:*
17: 16: sp_bacteriap:*
18: 17: sp_archaeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	72.7	1182	12 Q9ENL3	Q9enl3 colorado ti
2	45.5	68.9	269	10 Q9C9X6	Q9c9x6 arabidopsis
3	45.5	68.9	279	10 Q9LDE2	Q9lde2 arabidopsis
4	45.5	68.9	283	10 Q9SKD0	Q9skd0 arabidopsis
5	42	63.6	64	4 Q9UD19	Q9udi9 homo sapien
6	42	63.6	212	11 Q9Q929	Q9q929 mus musculu
7	42	63.6	287	4 Q96KT3	Q96kt3 homo sapien
8	42	63.6	332	4 Q99675	Q99675 homo sapien
9	42	63.6	332	4 Q96BX2	Q96bx2 homo sapien
10	42	63.6	332	11 P97587	P97587 rattus norv
11	42	63.6	342	5 Q16325	Q16325 caenorhabdi
12	42	63.6	347	5 Q16323	Q16323 caenorhabdi
13	42	63.6	351	5 Q16326	Q16326 caenorhabdi
14	42	63.6	2954	12 Q96898	Q96898 hepatitis g
15	42	63.6	2967	12 Q41892	Q41892 hepatitis g
16	41	62.1	883	16 Q9WYF1	Q9wyf1 thermotoga

17	41	62.1	126	5	Q9VNF1	Q9vnf1 drosophila
18	41	62.1	252	2	P95462	P95462 plectonema
19	41	62.1	261	5	Q9N5P2	Q9n5p2 caenorhabdi
20	41	62.1	280	10	Q9S7S3	Q9s7s3 arabidopsis
21	41	62.1	323	4	Q9H635	Q9h635 homo sapien
22	41	62.1	441	4	Q9NV55	Q9nv55 homo sapien
23	41	62.1	441	16	Q9VLRK0	Q9vlrk0 clostridium
24	41	62.1	2970	12	O56073	O56073 hepatitis g
25	40	60.6	109	11	Q9CZC9	Q9czc9 mus musculus
26	40	60.6	130	2	Q9R9T8	Q9r9t8 bacillus su
27	40	60.6	130	2	Q9L7W8	Q9l7w8 bacillus su
28	40	60.6	130	16	O34506	O34506 bacillus su
29	40	60.6	368	12	Q9LRT1	Q9lrt1 tupaia herp
30	40	60.6	396	10	O49647	O49647 arabidopsis
31	40	60.6	415	10	Q9LMJ5	Q9lmj5 arabidopsis
32	40	60.6	441	16	Q9KFW3	Q9kfw3 bacillus ha
33	40	60.6	456	3	P87172	P87172 schizosacch
34	40	60.6	499	2	Q9SD29	Q9sd29 streptomyce
35	40	60.6	520	16	Q9CM14	Q9cm14 pasteurella
36	40	60.6	575	5	Q95X29	Q95y29 caenorhabdi
37	40	60.6	630	5	O24222	O24222 drosophila
38	40	60.6	630	5	Q9VLJ6	Q9vlj6 drosophila
39	40	60.6	997	16	Q92LC4	Q92lc4 rhizobium m
40	39.5	59.8	53	11	O08633	O08633 mus musculus
41	39.5	59.8	62	6	Q28581	Q28581 ovis aries
42	39	59.1	53	6	Q29334	Q29334 sus scrofa
43	39	59.1	85	11	Q9JHY3	Q9jhy3 mus musculus
44	39	59.1	116	16	O4K911	O4kg11 bacillus ha
45	39	59.1	118	16	O31623	O31623 bacillus su

ALIGNMENTS

RESULT	1	
Q9ENL3		
ID	Q9ENL3	PRELIMINARY; PRT; 1182 AA.
AC	Q9ENL3;	
DT	01-MAR-2001	(TREMBlrel. 16, Created)
DT	01-MAR-2001	(TREMBlrel. 16, Last sequence update)
DT	01-DEC-2001	(TREMBlrel. 19, Last annotation update)

Query Match	72.7%	Score 48;	DB 12;	Length 1182;
Best Local Similarity	100.0%;	Pred. No. 28;		
Matches	6;	Conservative	0;	Mismatches
			0;	Indels
				0;
				Gaps
				0;

RESULT	2
Q9C9X6	
ID	Q9C9X6
AC	Q9C9X6;
DT	01-JUN-2001 (TRENBLrel, 17, Created)
	PRELIMINARY;
	PRT; 269 AA.

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE HYPOTHETICAL 30.4 KDA PROTEIN.
GN T23K23.3
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLOMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Miltschker J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana";
RL Nature 408:816-820(2000).
DR EMBL; AC012563; AAG52002.1; -.
KW Hypothetical protein.
SQ SEQUENCE 269 AA; 30392 MW; 2753AB3AD6063FE8 CRC64;

Query Match 68.9%; Score 45.5; DB 10; Length 269;
Best Local Similarity 77.8%; Pred. NO. 16;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 CY-WGCGYW 8
Db 190 CYRWGCGW 198
|| |||||
RESULT 3 PRELIMINARY; PRT; 279 AA.
O9LDE2
AC Q9LDE2
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE F10B6.5 (T5E21.17) (AT1G14680/F10B6.22).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Chao Q., Shinn P., Dunn P., Buehler E., Kahn S., Kim C., Walker M.,
RA Williams S., Altafi H., Araujo R., Conn L., Conway A.B., Gonzalez A.,
RA Hansen N.F., Huizar L., Kremenetskaia I., Lenz C., Li J., Liu S.,
RA Luros S., Rowley D., Schwartz J., Toriumi M., Vysotskaia V., Yu G.,
RA Davis R.W., Federspiel N.A., Theologis A., Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F10B6 from chromosome
I.";
RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RA Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

RN [3]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei Q., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsy N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.R.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Tambunga G., Altafi H., Bei Q., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsy N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Federspiel N.A.,
RA Theologis A., Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC T5E21 from chromosome
I.";
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006917; AAF79219.1; -.
DR EMBL; AC010657; AAF63172.1; -.
DR EMBL; AY058073; AAL24181.1; -.
SQ SEQUENCE 279 AA; 31168 MW; BAB996037E04D372 CRC64;

Query Match 68.9%; Score 45.5; DB 10; Length 279;
Best Local Similarity 77.8%; Pred. NO. 17;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 CY-WGCGYW 8
Db 200 CYRWGCGW 208
|| |||||
RESULT 4 PRELIMINARY; PRT; 283 AA.
O9SKD0
ID O9SKD0
AC O9SKD0
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE AT2G01930 PROTEIN.
GN AT2G01930.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLOMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,


```

RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RA thaliana.";
RA Nature 402:761-768(1999).
RA [2]
RA SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA;
RA Lin X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RA EMBL; AC006532; AAF18588.1; -.
RA SQ SEQUENCE 283 AA; 31648 MW; BAFBFCBD92262E07 CRC64;

Query Match 58.9%; Score 45.5; DB 10; Length 283;
Best Local Similarity 77.8%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 cywgcgw 8
Db 204 CYRGCWGW 212

RESULT 5
Q9UD19 ID Q9UD19 PRELIMINARY; PRT; 64 AA.
AC Q9UD19;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE ARYL-SULFATASE B (EC 3.1.6.1) (FRAGMENTS).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=93003385; PubMed=1390929;
RA Kobayashi T., Honke K., Jin T., Gasa S., Miyazaki T., Makita A.;
RA "Components and proteolytic processing sites of arylsulfatase B from
RA human placenta.";
RL Biochim. Biophys. Acta 1159:243-247(1992).
RT HSP; P15848; 1FSU.
FT NON_TER 1 1
FT NON_CONS 15 16
FT NON_CONS 46 47
FT NON_TER 64 64
FT NON_CONS 64 64
SQ SEQUENCE 64 AA; 7270 MW; B003FB1459CA8FB2 CRC64;

Query Match 63.6%; Score 42; DB 4; Length 64;
Best Local Similarity 85.7%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ywgcgyw 8
Db 36 YPGCGYW 42

RESULT 6
Q9D929 ID Q9D929 PRELIMINARY; PRT; 212 AA.
AC Q9D929;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 1810009H17RIK PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Saito R.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baidarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK007402; BAB25015.1; -.
DR MGP; MGI:1916368; 1810009H17RIK.
SQ SEQUENCE 212 AA; 24313 MW; 7D5EDF67496F61C4 CRC64;

Query Match 63.6%; Score 42; DB 11; Length 212;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cywgc 5
Db 98 CYWGC 102

RESULT 7
Q96KT3 ID Q96KT3 PRELIMINARY; PRT; 287 AA.
AC Q96KT3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 30.4 KDA PROTEIN (FRAGMENT).
GN C8ORF7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Appel S., Berghelm A., Reichwald K., Reis A., Rosenthal A., Ramsay M.,
RA Hennies H.;
RA "Transcript map of the KWE critical region on chromosome 8p22-p23.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ301560; CAC82736.1; -.
KW Hypothetical protein.
FT NON_TER 1 1
FT NON_TER 1 1
SQ SEQUENCE 287 AA; 30401 MW; 39234688A8B01A69 CRC64;

Query Match 63.6%; Score 42; DB 4; Length 287;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cywgc 5
Db 65 CYWGC 69

```

```
RESULT 8
Q99675 PRELIMINARY; PRT; 332 AA.
AC Q99675;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE CELL GROWTH REGULATOR CGR19.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=97122496; PubMed=8968090;
RA Madden S.L., Galella E.A., Riley D., Bertelsen A.H., Beaudry G.A.;
RT "Induction of cell growth regulatory genes by p53.";
RL Cancer Res. 56:5384-5390(1996).
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL: U66469; AAC50897.1; -.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF00097; ZF-C3HC4; 1.
DR SMART: SM00184; RING; 1.
KW zinc-finger.
SQ SEQUENCE 332 AA; 38241 MW; 2F1FCOD12B710C80 CRC64;
```

```
Query Match 63.6%; Score 42; DB 4; Length 332;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cywgc 5
| | | | |
DB 98 CYWGC 102
```

```
RESULT 9
Q96BX2 PRELIMINARY; PRT; 332 AA.
AC Q96BX2;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE CELL GROWTH REGULATOR WITH RING FINGER DOMAIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SKELETAL MUSCLE;
RA Strausberg R.;
RL Submitted (OCr-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC015063; AAI15063.1; -.
SQ SEQUENCE 332 AA; 38268 MW; 0FECCELEB87950FA CRC64;
```

```
Query Match 63.6%; Score 42; DB 4; Length 332;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cywgc 5
| | | | |
DB 98 CYWGC 102
```

```
RESULT 10
P97587 PRELIMINARY; PRT; 332 AA.
AC P97587;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
```

```
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE CELL GROWTH REGULATOR RGR19.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FISHER; TISSUE=FIBROBLAST;
RX MEDLINE=97122496; PubMed=8968090;
RA Madden S.L., Galella E.A., Riley D., Bertelsen A.H., Beaudry G.A.;
RT "Induction of cell growth regulatory genes by p53.";
RL Cancer Res. 56:5384-5390(1996).
DR EMBL: U66471; AAC52951.1; -.
DR InterPro: IPR001841; Znf_ring.
DR SMART: SM00184; RING; 1.
DR SNAPT: SM00184; RING; 1.
SQ SEQUENCE 332 AA; 37443 MW; D918EC9B74885104 CRC64;
```

```
Query Match 63.6%; Score 42; DB 11; Length 332;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cywgc 5
| | | | |
DB 98 CYWGC 102
```

```
RESULT 11
O16325 PRELIMINARY; PRT; 342 AA.
AC O16325;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 38.5 KDA PROTEIN.
GN C07G3.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX Geisel C., Wamsley P.;
RT "The sequence of C. elegans cosmid C07G3.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF016432; AAB65382.1; -.
DR InterPro: IPR003002; 7TM_chemorecept_1.
DR Pfam: PF01461; 7tm_4; 1.
KW Hypothetical protein.
SQ SEQUENCE 342 AA; 38469 MW; 6EF773489493954D CRC64;
```

```
Query Match 63.6%; Score 42; DB 5; Length 342;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_RdRP.
DR InterPro: IPR003880; Phosphopant_attach.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR Pfam: PF01006; HCV_NS4a; 1.
DR Pfam: PF01001; HCV_NS4b; 1.
DR Pfam: PF01506; HCV_NS5a; 1.
DR Pfam: PF00998; HCV_RdRP; 1.
DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
SQ SEQUENCE 2954 AA; 318833 MW; 4A0C35FC2ED283B1 CRC64;

Query Match 63.6%; Score 42; DB 12; Length 2954;
Best Local Similarity 62.5%; Pred. No. 5e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 cywgcgyw 8
   | | | |
Db 277 CQWGSAYW 284

RESULT 15
O41892
ID O41892 PRELIMINARY; PRT; 2967 AA.
AC O41892;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GENOME POLYPROTEIN.
OS Hepatitis GB virus A.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae.
OX NCBI_TaxID=39112;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ALAB;
RX MEDLINE=971126113; PubMed=8971037;
RA Leary T.P., Desai S.M., Yamaguchi J., Chalmers M.L., Schlauder G.G.,
RA Dawson G.J., Mushahwar I.K.;
RT "Species-specific variants of GB virus A in captive monkeys.";
RL J. Virol. 70:9028-9030(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ALAB;
RX MEDLINE=97437486; PubMed=9292019;
RA Leary T.P., Desai S.M., Erker J.C., Mushahwar I.K.;
RT "The sequence and genomic organization of a GB virus A variant
isolated from captive tamarins.";
RL J. Gen. Virol. 78:0-0(0).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ALAB;
RA Leary T.P., Desai S.M., Erker J.C., Mushahwar I.K.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBDJ databases.
DR EMBL: U94421; AAB71133.1; -.
DR HSSP: P27958; 1HEI.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002518; HCV_NS2.
DR InterPro: IPR004109; HCV_NS3.
DR InterPro: IPR000745; HCV_NS4a.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_RdRP.
DR InterPro: IPR000366; Lipocin_cytfabp.
DR InterPro: IPR001005; Myb_DNA_bind.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR Pfam: PF01006; HCV_NS4a; 1.
DR Pfam: PF01001; HCV_NS4b; 1.
DR Pfam: PF01506; HCV_NS5a; 1.
DR Pfam: PF00998; HCV_RdRP; 1.
DR PROSITE: PS00013; LIPOCALIN; UNKNOWN_1.

DR PROSITE: PS00037; MYB_1; UNKNOWN_1.
KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
SQ SEQUENCE 2967 AA; 321140 MW; 9AAB97B830C5D199 CRC64;

Query Match 63.6%; Score 42; DB 12; Length 2967;
Best Local Similarity 62.5%; Pred. No. 5e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 cywgcgyw 8
   | | | |
Db 277 CQWGSAYW 284

Search completed: August 12, 2002, 17:28:16
Job time: 363 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:21:41 ; Search time 51.76 seconds
(without alignments)
17.168 Million cell updates/sec

Title: 10-071247-2

Perfect score: 66

Sequence: 1 cywgcgyw 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

- 1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	66	100.0	8	20	AA1980
2	66	100.0	8	21	AA1981
3	48	72.7	54	21	AA1982
4	47	71.2	841	21	AA1983
5	45.5	68.9	233	21	AA1984
6	45.5	68.9	238	21	AA1985
7	45.5	68.9	252	21	AA1986
8	45.5	68.9	252	21	AA1987
9	45.5	68.9	252	21	AA1988
10	45.5	68.9	252	21	AA1989
11	45.5	68.9	279	21	AA1990

12	45.5	68.9	279	21	AA1991
13	45.5	68.9	279	21	AA1992
14	45.5	68.9	282	21	AA1993
15	45.5	68.9	285	21	AA1994
16	45.5	68.9	286	21	AA1995
17	45	68.2	8	20	AA1996
18	45	68.2	8	20	AA1997
19	45	68.2	9	21	AA1998
20	45	68.2	75	22	AA1999
21	43	65.2	60	22	AA2000
22	42	63.6	7	20	AA2001
23	42	63.6	115	21	AA2002
24	42	63.6	332	19	AA2003
25	42	63.6	332	19	AA2004
26	42	63.6	332	20	AA2005
27	42	63.6	345	22	AA2006
28	42	63.6	533	22	AA2007
29	42	63.6	2972	21	AA2008
30	42	63.6	3163	16	AA2009
31	41	62.1	32	22	AA2010
32	41	62.1	32	22	AA2011
33	41	62.1	32	22	AA2012
34	41	62.1	32	22	AA2013
35	41	62.1	69	22	AA2014
36	41	62.1	76	22	AA2015
37	41	62.1	126	22	AA2016
38	41	62.1	358	10	AA2017
39	41	62.1	441	22	AA2018
40	41	62.1	602	22	AA2019
41	41	62.1	1851	22	AA2020
42	40.5	61.4	1604	22	AA2021
43	40	60.6	7	20	AA2022
44	40	60.6	105	22	AA2023
45	40	60.6	125	21	AA2024

ALIGNMENTS

RESULT 1

AA1980 15
ID AA1980 standard; peptide: 8 AA.

AC AA1980;

DT 08-JUN-1999 (first entry)

DE Fluorine-18 (F-18) labeled peptide 2.

18F radionuclide; targeting vector; positron emission tomography; F-18;
radiolabeling; thiol; fluorine-18.

OS Synthetic.

Key Location/Qualifiers

Misc-difference 1 /note= "N-terminal acetylation; optionally has a free
or protected thiol group"

Misc-difference 2 /note= "D-form residue"

Misc-difference 3 /note= "D-form residue"

Misc-difference 5 /note= "D-form residue"

Misc-difference 7 /note= "D-form residue; optionally has a free or
protected thiol group"

Misc-difference 8 /note= "D-form residue"

Misc-difference /note= "D-form residue"

WO9911590-A1.

11-MAR-1999.


```

XX PN WO9958552-A2.
XX PD
XX PF 18-NOV-1999.
XX XX
XX PF 03-MAY-1999; 99WO-NO00143.
XX XX
XX PR 08-MAY-1998; 98NO-0002097.
XX XX
XX PA (NHSD ) NORSK HYDRO AS.
XX XX
XX PI Gaudernack G, Eriksen JA, Moller M, Gjertsen MK, Saeterdal I;
XX XX WPI; 2000-039064/03.
XX DR
XX XX
XX PT New peptides derived from genes with frameshift mutations, used to
XX XX develop products for the treatment and prophylaxis of cancers -
XX PS Claim 13; Page 36; 166pp; English.
XX XX
XX CC Peptides AAY65684-Y66142 are fragments of mutant proteins arising from a
XX CC frameshift mutation in a gene from a cancer cell. The peptides are
XX CC characterised in that they:
XX CC (i) are at least 8 amino acids long and a fragment of a mutant protein
XX CC arising from a frameshift mutation in a gene of a cancer cell;
XX CC (ii) consist of at least one amino acid of the mutant part of a protein
XX CC sequence encoded by the gene;
XX CC (iii) comprise 0-10 amino acid from the carboxyl terminus of the normal
XX CC part of the protein sequence preceding the amino terminus of the mutant
XX CC sequence and may further extend to the carboxyl terminus of the mutant
XX CC part of the protein as determined by a new stop codon generated by the
XX CC frameshift mutation; and
XX CC (iv) induce, either in their full lengths or after processing by an
XX CC antigen presenting cell (APC), T cell responses.
XX CC The genes that the peptides are derived from, are characterised as
XX CC susceptible to frameshift mutation by having a mono nucleoside base
XX CC repeat sequence of at least 5 residues, or a di-nucleoside base repeat
XX CC sequence of at least 4 di-nucleoside base units. The peptides are
XX CC created by the addition or deletion of 1 or 2 nucleoside base residues
XX CC from the repeat sequence. The novel peptides can elicit T cell responses
XX CC and toxicity against tumours and cancer cells carrying genes with
XX CC frameshift mutations. The novel peptides and DNA sequences can be used
XX CC for the preparation of a composition for the treatment or prophylaxis of
XX CC cancer.
XX XX
XX SQ Sequence 54 AA;

Query Match 72.7%; Score 48; DB 21; Length 54;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 cywgcgww 8
Db 5 cewgcgsw 12

RESULT 4
AAB18511
ID AAB18511 standard; protein; 841 AA.
AC AAB18511;
XX XX
XX DT 15-JAN-2001 (first entry)
XX XX
XX DE H2 homologue of prolyl-tripeptidyl peptidase DPP.
XX XX
XX KW Prolyl tripeptidyl-peptidase; amidolytic activity; periodontal disease;
XX XX gingivitis; periodontitis.
XX OS Porphyromonas gingivalis.
XX XX
XX PN WO200052147-A2.

```

```

XX PD 08-SEP-2000.
XX XX
XX PF 03-MAR-2000; 2000WO-US05551.
XX XX
XX PR 05-MAR-1999; 99US-0123148.
XX XX
XX PA (UYGE-) UNIV GEORGIA RES FOUND INC.
XX PA (TRAV/) TRAVIS J.
XX PA (POTE/) POTEMPA J.
XX PA (BANB/) BANBULA A.
XX XX
XX PI Travis J, Potempa J, Banbula A;
XX XX WPI; 2000-594181/56.
XX DR
XX XX
XX PT prolyl tripeptidyl-peptidase, active analog, fragment or variant useful
XX XX for identifying its inhibitor which is useful for protecting an animal
XX XX from a periodontal disease such as gingivitis and periodontitis -
XX PS Claim 22; Fig 6; 58pp; English.
XX XX
XX CC The present sequence represents a H2 homologue of a prolyl
XX CC tripeptidyl-peptidase (designated DPP) from Porphyromonas gingivalis.
XX CC The prolyl tripeptidyl-peptidase has an amidolytic activity, and cleaves
XX CC a peptide bond in a target polypeptide having at least 4 amino acids.
XX CC This bond is between a proline and an amino acid attached to the
XX CC alpha-carboxyl group end of the proline. The polypeptide is useful
XX CC for identifying inhibitors. These inhibitors are then useful for
XX CC reducing the growth of bacterium or for protecting an animal from a
XX CC periodontal disease such as gingivitis and periodontitis caused by
XX CC Porphyromonas gingivalis.
XX XX
XX SQ Sequence 841 AA;

Query Match 71.2%; Score 47; DB 21; Length 841;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywgcgww 8
Db 719 ywgsqww 725

RESULT 5
AAG40353
ID AAG40353 standard; Protein; 233 AA.
XX XX
XX AC AAG40353;
XX XX
XX DT 18-OCT-2000 (first entry)
XX XX
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 50058.
XX XX
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX XX
XX OS Arabidopsis thaliana.
XX XX
XX PN EP1033405-A2.
XX XX
XX PD 06-SEP-2000.
XX XX
XX XX
XX XX 25-FEB-2000; 2000EP-0301439.
XX XX
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.

```

PR 01-APR-1999;	99US-0127462.	PR 19-JUL-1999;	99US-0144331.
PR 06-APR-1999;	99US-0128234.	PR 19-JUL-1999;	99US-0144332.
PR 08-APR-1999;	99US-0128714.	PR 19-JUL-1999;	99US-0144333.
PR 16-APR-1999;	99US-0129845.	PR 19-JUL-1999;	99US-0144334.
PR 19-APR-1999;	99US-0130077.	PR 19-JUL-1999;	99US-0144335.
PR 21-APR-1999;	99US-0130449.	PR 20-JUL-1999;	99US-0144352.
PR 23-APR-1999;	99US-0130510.	PR 20-JUL-1999;	99US-0144632.
PR 23-APR-1999;	99US-0130891.	PR 20-JUL-1999;	99US-0144884.
PR 28-APR-1999;	99US-0131449.	PR 21-JUL-1999;	99US-0144814.
PR 30-APR-1999;	99US-0132048.	PR 21-JUL-1999;	99US-0145086.
PR 30-APR-1999;	99US-0132407.	PR 21-JUL-1999;	99US-0145088.
PR 04-MAY-1999;	99US-0132484.	PR 22-JUL-1999;	99US-0145085.
PR 05-MAY-1999;	99US-0132485.	PR 22-JUL-1999;	99US-0145087.
PR 06-MAY-1999;	99US-0132486.	PR 22-JUL-1999;	99US-0145089.
PR 07-MAY-1999;	99US-0132487.	PR 22-JUL-1999;	99US-0145192.
PR 11-MAY-1999;	99US-0132863.	PR 23-JUL-1999;	99US-0145145.
PR 11-MAY-1999;	99US-0134256.	PR 23-JUL-1999;	99US-0145218.
PR 14-MAY-1999;	99US-0134218.	PR 23-JUL-1999;	99US-0145224.
PR 14-MAY-1999;	99US-0134219.	PR 26-JUL-1999;	99US-0145276.
PR 14-MAY-1999;	99US-0134221.	PR 27-JUL-1999;	99US-0145913.
PR 14-MAY-1999;	99US-0134370.	PR 27-JUL-1999;	99US-0145918.
PR 18-MAY-1999;	99US-0134768.	PR 27-JUL-1999;	99US-0145919.
PR 19-MAY-1999;	99US-0134941.	PR 28-JUL-1999;	99US-0145951.
PR 20-MAY-1999;	99US-0135124.	PR 02-AUG-1999;	99US-0146386.
PR 21-MAY-1999;	99US-0135353.	PR 02-AUG-1999;	99US-0148388.
PR 24-MAY-1999;	99US-0135629.	PR 02-AUG-1999;	99US-0148389.
PR 25-MAY-1999;	99US-0136021.	PR 03-AUG-1999;	99US-0147038.
PR 27-MAY-1999;	99US-0136392.	PR 04-AUG-1999;	99US-0147204.
PR 28-MAY-1999;	99US-0136782.	PR 04-AUG-1999;	99US-0147302.
PR 01-JUN-1999;	99US-0137222.	PR 05-AUG-1999;	99US-0147192.
PR 03-JUN-1999;	99US-0137528.	PR 05-AUG-1999;	99US-0147260.
PR 04-JUN-1999;	99US-0137502.	PR 06-AUG-1999;	99US-0147303.
PR 07-JUN-1999;	99US-0137724.	PR 06-AUG-1999;	99US-0147416.
PR 08-JUN-1999;	99US-0138094.	PR 09-AUG-1999;	99US-0147493.
PR 10-JUN-1999;	99US-0138540.	PR 09-AUG-1999;	99US-0147935.
PR 10-JUN-1999;	99US-0138847.	PR 10-AUG-1999;	99US-0148171.
PR 14-JUN-1999;	99US-0139119.	PR 11-AUG-1999;	99US-0148319.
PR 16-JUN-1999;	99US-0139432.	PR 12-AUG-1999;	99US-0148341.
PR 16-JUN-1999;	99US-0139453.	PR 13-AUG-1999;	99US-0148565.
PR 17-JUN-1999;	99US-0139492.	PR 13-AUG-1999;	99US-0148684.
PR 18-JUN-1999;	99US-0139454.	PR 16-AUG-1999;	99US-0149368.
PR 18-JUN-1999;	99US-0139455.	PR 17-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139456.	PR 18-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139457.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139459.	PR 20-AUG-1999;	99US-0149929.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139461.	PR 23-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139462.	PR 25-AUG-1999;	99US-0150566.
PR 18-JUN-1999;	99US-0139463.	PR 26-AUG-1999;	99US-0150884.
PR 18-JUN-1999;	99US-0139750.	PR 27-AUG-1999;	99US-0151065.
PR 18-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151066.
PR 21-JUN-1999;	99US-0139817.	PR 27-AUG-1999;	99US-0151080.
PR 22-JUN-1999;	99US-0139899.	PR 30-AUG-1999;	99US-0151303.
PR 23-JUN-1999;	99US-0140353.	PR 31-AUG-1999;	99US-0151438.
PR 23-JUN-1999;	99US-0140354.	PR 01-SEP-1999;	99US-0151920.
PR 24-JUN-1999;	99US-0140695.	PR 07-SEP-1999;	99US-0152363.
PR 28-JUN-1999;	99US-0140823.	PR 10-SEP-1999;	99US-0153070.
PR 29-JUN-1999;	99US-0140991.	PR 13-SEP-1999;	99US-0153758.
PR 30-JUN-1999;	99US-0141287.	PR 15-SEP-1999;	99US-0154018.
PR 01-JUL-1999;	99US-0141842.	PR 16-SEP-1999;	99US-0154039.
PR 01-JUL-1999;	99US-0142154.	PR 20-SEP-1999;	99US-0154779.
PR 02-JUL-1999;	99US-0142053.	PR 22-SEP-1999;	99US-0155139.
PR 06-JUL-1999;	99US-0142390.	PR 23-SEP-1999;	99US-0155486.
PR 08-JUL-1999;	99US-0142803.	PR 24-SEP-1999;	99US-0155659.
PR 09-JUL-1999;	99US-0142920.	PR 28-SEP-1999;	99US-0156458.
PR 12-JUL-1999;	99US-0142977.	PR 29-SEP-1999;	99US-0156596.
PR 13-JUL-1999;	99US-0143542.	PR 04-OCT-1999;	99US-0157117.
PR 14-JUL-1999;	99US-0143624.	PR 05-OCT-1999;	99US-0157753.
PR 15-JUL-1999;	99US-0144005.	PR 08-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.


```
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      58.9%; Score 45.5; DB 21; Length 233;
Best Local Similarity 77.8%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Caps 1;

QY 1 cy-wgcgww 8
   |||||
Db 154 cyrwgcgww 162

RESULT 6
AAG60115
ID AAG60115 standard; Protein; 238 AA.
XX
AC AAG60115;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 77829.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
```

```
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158023.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      68.9%; Score 45.5; DB 21; Length 238;
Best Local Similarity 77.8%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 cy-wgcgww 8
   || |||||
Db 159 cyrwgcgw 167

RESULT 7
AAG05644
ID AAG05644 standard; Protein; 252 AA.
XX
AC AAG05644;
XX
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 2124.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
```

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PR	21-OCT-1999;	99US-0160815.	PR	18-MAY-1999;	99US-0134768.
PR	22-OCT-1999;	99US-0160980.	PR	19-MAY-1999;	99US-0134941.
PR	22-OCT-1999;	99US-0160981.	PR	20-MAY-1999;	99US-0135124.
PR	22-OCT-1999;	99US-0160989.	PR	21-MAY-1999;	99US-0135353.
PR	25-OCT-1999;	99US-0161404.	PR	24-MAY-1999;	99US-0135629.
PR	25-OCT-1999;	99US-0161409.	PR	25-MAY-1999;	99US-0136021.
PR	25-OCT-1999;	99US-0161406.	PR	27-MAY-1999;	99US-0136392.
PR	26-OCT-1999;	99US-0161359.	PR	28-MAY-1999;	99US-0136782.
PR	26-OCT-1999;	99US-0161360.	PR	01-JUN-1999;	99US-0137222.
PR	26-OCT-1999;	99US-0161361.	PR	03-JUN-1999;	99US-0137528.
PR	28-OCT-1999;	99US-0161920.	PR	04-JUN-1999;	99US-0137502.
PR	28-OCT-1999;	99US-0161992.	PR	07-JUN-1999;	99US-0137724.
PR	28-OCT-1999;	99US-0161993.	PR	08-JUN-1999;	99US-0138094.
PR	29-OCT-1999;	99US-0162142.	PR	10-JUN-1999;	99US-0138540.
Query Match 68.9%; Score 45.5; DB 21; Length 252;					
Best Local Similarity 77.8%; Pred. No. 86;					
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;					
QY	1 cy-wgcgww 8				
Db	173 cyrwgcgww 181				
RESULT 8					
AA	AAG38268				
ID	AAG38268 standard; Protein; 252 AA.				
XX					
AC	AAG38268;				
XX					
DT	18-OCT-2000 (first entry)				
XX					
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 47184.				
XX					
KW	Protein identification; signal transduction pathway; metabolic pathway;				
KW	hybridisation assay; genetic mapping; gene expression control; promoter;				
KW	termination sequence.				
XX					
OS	Arabidopsis thaliana.				
XX					
PN	EP1033405-A2.				
XX					
PD	06-SEP-2000.				
XX					
PF	25-FEB-2000; 2000EP-0301439.				
XX					
PR	25-FEB-1999;	99US-0121825.	PR	18-MAY-1999;	99US-0134768.
PR	05-MAR-1999;	99US-0123180.	PR	19-MAY-1999;	99US-0134941.
PR	09-MAR-1999;	99US-0123548.	PR	20-MAY-1999;	99US-0135124.
PR	23-MAR-1999;	99US-0125788.	PR	21-MAY-1999;	99US-0135353.
PR	25-MAR-1999;	99US-0126264.	PR	24-MAY-1999;	99US-0135629.
PR	29-MAR-1999;	99US-0126785.	PR	25-MAY-1999;	99US-0136021.
PR	01-APR-1999;	99US-0127462.	PR	27-MAY-1999;	99US-0136392.
PR	06-APR-1999;	99US-0128234.	PR	28-MAY-1999;	99US-0136782.
PR	08-APR-1999;	99US-0128714.	PR	01-JUN-1999;	99US-0137222.
PR	16-APR-1999;	99US-0129845.	PR	03-JUN-1999;	99US-0137528.
PR	19-APR-1999;	99US-0130077.	PR	04-JUN-1999;	99US-0137502.
PR	21-APR-1999;	99US-0130449.	PR	07-JUN-1999;	99US-0137724.
PR	23-APR-1999;	99US-0130510.	PR	08-JUN-1999;	99US-0138094.
PR	28-APR-1999;	99US-0130891.	PR	10-JUN-1999;	99US-0138540.
PR	30-APR-1999;	99US-0131449.	PR	10-JUN-1999;	99US-0138847.
PR	30-APR-1999;	99US-0132048.	PR	14-JUN-1999;	99US-0139119.
PR	04-MAY-1999;	99US-0132407.	PR	16-JUN-1999;	99US-0139452.
PR	05-MAY-1999;	99US-0132484.	PR	16-JUN-1999;	99US-0139453.
PR	06-MAY-1999;	99US-0132485.	PR	17-JUN-1999;	99US-0139492.
PR	06-MAY-1999;	99US-0132486.	PR	18-JUN-1999;	99US-0139454.
PR	07-MAY-1999;	99US-0132487.	PR	18-JUN-1999;	99US-0139455.
PR	11-MAY-1999;	99US-0132863.	PR	18-JUN-1999;	99US-0139456.
PR	14-MAY-1999;	99US-0134256.	PR	18-JUN-1999;	99US-0139457.
PR	14-MAY-1999;	99US-0134218.	PR	18-JUN-1999;	99US-0139458.
PR	14-MAY-1999;	99US-0134219.	PR	18-JUN-1999;	99US-0139459.
PR	14-MAY-1999;	99US-0134221.	PR	18-JUN-1999;	99US-0139460.
PR	14-MAY-1999;	99US-0134370.	PR	18-JUN-1999;	99US-0139461.

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 68.98; Score 45.5; DB 21; Length 252;
Best Local Similarity 77.8%;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 cy-wgcgww 8
|| |||||
Db 173 cyrwgcgww 181

RESULT 9
AAG40352
ID AAG40352 standard; Protein; 252 AA.
XX
AC AAG40352;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 50057.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.

PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	17-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139463.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140033.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140981.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161992.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0162142.

Query Match	68.9%;	Score 45.5;	DB 21;	Length 252;
Best Local Similarity	77.8%;	Pred. No. 86;		
Matches 7;	Conservative 0;	Mismatches 1;	Indels 1;	Gaps 1;
Qy	1 cy-wgcgwyw 8			
Db	173 cyrwcgwgw 181			
RESULT 10				
AAG60114	ID ID1	AAG60114 standard; Protein; 257 AA.		
XX	XX			
AC	AAG60114;			
DT	18-OCT-2000 (first entry)			
XX	XX			
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 77828.			
XX	XX			
KW	Protein identification; signal transduction pathway; metabolic pathway;			
KW	hybridisation assay; genetic mapping; gene expression control; promoter;			
KW	termination sequence.			
XX	XX			
OS	Arabidopsis thaliana.			
XX	XX			
PN	EPI033405-A2.			
XX	XX			
PD	06-SEP-2000.			
PF	25-FEB-2000; 2000EP-0301439.			
XX	XX			
PR	25-FEB-1999; 99US-0121825.			
PR	05-MAR-1999; 99US-0123180.			
PR	09-MAR-1999; 99US-0123548.			
PR	23-MAR-1999; 99US-0125788.			
PR	25-MAR-1999; 99US-0126264.			
PR	29-MAR-1999; 99US-0126785.			
PR	01-APR-1999; 99US-0127452.			
PR	06-APR-1999; 99US-0128234.			
PR	08-APR-1999; 99US-0128714.			
PR	16-APR-1999; 99US-0129845.			
PR	19-APR-1999; 99US-0130077.			
PR	21-APR-1999; 99US-0130449.			
PR	23-APR-1999; 99US-0130510.			
PR	28-APR-1999; 99US-0130891.			
PR	30-APR-1999; 99US-0131449.			
PR	30-APR-1999; 99US-0132048.			
PR	04-MAY-1999; 99US-0132407.			
PR	05-MAY-1999; 99US-0132484.			
PR	06-MAY-1999; 99US-0132485.			
PR	06-MAY-1999; 99US-0132486.			
PR	07-MAY-1999; 99US-0132487.			
PR	11-MAY-1999; 99US-0132863.			
PR	14-MAY-1999; 99US-0134256.			
PR	14-MAY-1999; 99US-0134218.			
PR	14-MAY-1999; 99US-0134219.			
PR	14-MAY-1999; 99US-0134221.			
PR	18-MAY-1999; 99US-0134370.			
PR	19-MAY-1999; 99US-0134768.			
PR	19-MAY-1999; 99US-0134941.			
PR	20-MAY-1999; 99US-0135124.			
PR	21-MAY-1999; 99US-0135353.			
PR	24-MAY-1999; 99US-0135629.			
PR	25-MAY-1999; 99US-0136021.			
PR	27-MAY-1999; 99US-0136392.			
PR	28-MAY-1999; 99US-0136782.			
PR	01-JUN-1999; 99US-0137222.			
PR	03-JUN-1999; 99US-0137528.			
PR	04-JUN-1999; 99US-0137502.			
PR	07-JUN-1999; 99US-0137724.			
PR	08-JUN-1999; 99US-0138094.			
PR	10-JUN-1999; 99US-0138340.			
PR	10-JUN-1999; 99US-0138847.			
PR	14-JUN-1999; 99US-0139119.			
PR	16-JUN-1999; 99US-0139452.			
PR	16-JUN-1999; 99US-0139453.			
PR	17-JUN-1999; 99US-0139492.			
PR	18-JUN-1999; 99US-0139454.			
PR	18-JUN-1999; 99US-0139455.			
PR	18-JUN-1999; 99US-0139456.			
PR	18-JUN-1999; 99US-0139457.			
PR	18-JUN-1999; 99US-0139458.			
PR	18-JUN-1999; 99US-0139459.			
PR	18-JUN-1999; 99US-0139460.			
PR	18-JUN-1999; 99US-0139461.			
PR	18-JUN-1999; 99US-0139462.			
PR	18-JUN-1999; 99US-0139463.			
PR	18-JUN-1999; 99US-0139750.			
PR	18-JUN-1999; 99US-0139763.			
PR	21-JUN-1999; 99US-0139817.			
PR	22-JUN-1999; 99US-0139899.			
PR	23-JUN-1999; 99US-0140353.			
PR	23-JUN-1999; 99US-0140354.			
PR	24-JUN-1999; 99US-0140695.			
PR	28-JUN-1999; 99US-0140823.			
PR	29-JUN-1999; 99US-0140991.			

PR 10-AUG-1999;	99US-01481171.	Db 178 cyrwgcggw 186
PR 11-AUG-1999;	99US-0148319.	
PR 12-AUG-1999;	99US-0148341.	
PR 13-AUG-1999;	99US-0148565.	
PR 13-AUG-1999;	99US-0148684.	
PR 16-AUG-1999;	99US-0149368.	
PR 17-AUG-1999;	99US-0149175.	
PR 18-AUG-1999;	99US-0149426.	
PR 20-AUG-1999;	99US-0149722.	
PR 20-AUG-1999;	99US-0149723.	
PR 20-AUG-1999;	99US-0149929.	
PR 23-AUG-1999;	99US-0149902.	
PR 23-AUG-1999;	99US-0149930.	
PR 25-AUG-1999;	99US-0150566.	
PR 26-AUG-1999;	99US-0150884.	
PR 27-AUG-1999;	99US-0151065.	
PR 27-AUG-1999;	99US-0151066.	
PR 27-AUG-1999;	99US-0151080.	
PR 30-AUG-1999;	99US-0151303.	
PR 31-AUG-1999;	99US-0151438.	
PR 01-SEP-1999;	99US-0151930.	
PR 07-SEP-1999;	99US-0152363.	
PR 10-SEP-1999;	99US-0153070.	
PR 13-SEP-1999;	99US-0153758.	
PR 15-SEP-1999;	99US-0154018.	
PR 16-SEP-1999;	99US-0154039.	
PR 20-SEP-1999;	99US-0154779.	
PR 22-SEP-1999;	99US-0155139.	
PR 23-SEP-1999;	99US-0155148.	
PR 24-SEP-1999;	99US-0155659.	
PR 28-SEP-1999;	99US-0156458.	
PR 29-SEP-1999;	99US-0156596.	
PR 04-OCT-1999;	99US-0157117.	
PR 05-OCT-1999;	99US-0157753.	
PR 06-OCT-1999;	99US-0157865.	
PR 07-OCT-1999;	99US-0158029.	
PR 08-OCT-1999;	99US-0158232.	
PR 12-OCT-1999;	99US-0158369.	
PR 13-OCT-1999;	99US-0159293.	
PR 13-OCT-1999;	99US-0159294.	
PR 13-OCT-1999;	99US-0159295.	
PR 14-OCT-1999;	99US-0159329.	
PR 14-OCT-1999;	99US-0159330.	
PR 14-OCT-1999;	99US-0159331.	
PR 14-OCT-1999;	99US-0159637.	
PR 14-OCT-1999;	99US-0159638.	
PR 18-OCT-1999;	99US-0159584.	
PR 21-OCT-1999;	99US-0160741.	
PR 21-OCT-1999;	99US-0160767.	
PR 21-OCT-1999;	99US-0160768.	
PR 21-OCT-1999;	99US-0160770.	
PR 21-OCT-1999;	99US-0160814.	
PR 21-OCT-1999;	99US-0160815.	
PR 22-OCT-1999;	99US-0160980.	
PR 22-OCT-1999;	99US-0160981.	
PR 22-OCT-1999;	99US-0160989.	
PR 25-OCT-1999;	99US-0161404.	
PR 25-OCT-1999;	99US-0161405.	
PR 25-OCT-1999;	99US-0161406.	
PR 26-OCT-1999;	99US-0161359.	
PR 26-OCT-1999;	99US-0161360.	
PR 26-OCT-1999;	99US-0161361.	
PR 28-OCT-1999;	99US-0161920.	
PR 28-OCT-1999;	99US-0161992.	
PR 28-OCT-1999;	99US-0161993.	
PR 29-OCT-1999;	99US-0162142.	
<div>Query Match 68.9%; Score 45.5; DB 21; Length 257; Best Local Similarity 77.8%; Pred. No. 87; Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;</div>		
<div>QY 1 cy-wccgyw 8 </div>		

RESULT 11

AG05643 standard; Protein; 279 AA.

AC AAG05643;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 2123.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 23-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144634.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.

PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 68.9%; Score 45.5; DB 21; Length 279;
Best Local Similarity 77.8%; Pred. No. 93;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 cy-wgcgyw 8
|||
Db 200 cyrwgcgwg 208

RESULT 12
AAG38267
ID AAG38267 standard; Protein; 279 AA.
XX


```
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 68.9%; Score 45.5; DB 21; Length 279;
Best Local Similarity 77.8%; Pred. No. 93;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 cy-wgcqyw 8
Db 200 cyrwgcgw 208
|| |||||

RESULT 13
ID AAG40351 standard; Protein; 279 AA.
XX AC AAG40351;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 50056.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
```

```
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
```

PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.

PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 68.9%; Score 45.5; DB 21; Length 279;
Best Local Similarity 77.8%; Pred. No. 93;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 cy-wgcgww 8
|| |||||
Db 200 cywgcgww 208

RESULT 14

AAG38266
ID AAG38266 standard; Protein; 282 AA.
XX
AC AAG38266;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47182.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX

PR	23-SEP-1999;	99US-0155486.	PR	23-MAR-1999;	99US-0125788.
PR	24-SEP-1999;	99US-0155659.	PR	25-MAR-1999;	99US-0126264.
PR	28-SEP-1999;	99US-0156458.	PR	29-MAR-1999;	99US-0126785.
PR	29-SEP-1999;	99US-0156596.	PR	01-APR-1999;	99US-0127462.
PR	04-OCT-1999;	99US-0157717.	PR	06-APR-1999;	99US-0128234.
PR	05-OCT-1999;	99US-0157753.	PR	08-APR-1999;	99US-0128714.
PR	06-OCT-1999;	99US-0157865.	PR	16-APR-1999;	99US-0129845.
PR	07-OCT-1999;	99US-0158029.	PR	19-APR-1999;	99US-0130077.
PR	08-OCT-1999;	99US-0158232.	PR	21-APR-1999;	99US-0130449.
PR	12-OCT-1999;	99US-0158369.	PR	23-APR-1999;	99US-0130510.
PR	13-OCT-1999;	99US-0159293.	PR	23-APR-1999;	99US-0130891.
PR	13-OCT-1999;	99US-0159294.	PR	28-APR-1999;	99US-0131449.
PR	13-OCT-1999;	99US-0159295.	PR	30-APR-1999;	99US-0132048.
PR	14-OCT-1999;	99US-0159329.	PR	30-APR-1999;	99US-0132407.
PR	14-OCT-1999;	99US-0159330.	PR	04-MAY-1999;	99US-0132484.
PR	14-OCT-1999;	99US-0159331.	PR	05-MAY-1999;	99US-0132485.
PR	14-OCT-1999;	99US-0159637.	PR	06-MAY-1999;	99US-0132486.
PR	14-OCT-1999;	99US-0159638.	PR	06-MAY-1999;	99US-0132487.
PR	18-OCT-1999;	99US-0159584.	PR	07-MAY-1999;	99US-0132863.
PR	21-OCT-1999;	99US-0160741.	PR	11-MAY-1999;	99US-0134256.
PR	21-OCT-1999;	99US-0160767.	PR	14-MAY-1999;	99US-0134218.
PR	21-OCT-1999;	99US-0160768.	PR	14-MAY-1999;	99US-0134219.
PR	21-OCT-1999;	99US-0160770.	PR	14-MAY-1999;	99US-0134221.
PR	21-OCT-1999;	99US-0160814.	PR	14-MAY-1999;	99US-0134370.
PR	21-OCT-1999;	99US-0160815.	PR	18-MAY-1999;	99US-0134768.
PR	22-OCT-1999;	99US-0160980.	PR	19-MAY-1999;	99US-0134941.
PR	22-OCT-1999;	99US-0160981.	PR	20-MAY-1999;	99US-0135124.
PR	22-OCT-1999;	99US-0160989.	PR	21-MAY-1999;	99US-0135353.
PR	22-OCT-1999;	99US-0161040.	PR	24-MAY-1999;	99US-0135629.
PR	25-OCT-1999;	99US-0161405.	PR	25-MAY-1999;	99US-0136021.
PR	25-OCT-1999;	99US-0161406.	PR	27-MAY-1999;	99US-0136392.
PR	26-OCT-1999;	99US-0161359.	PR	28-MAY-1999;	99US-0136782.
PR	26-OCT-1999;	99US-0161360.	PR	01-JUN-1999;	99US-0137222.
PR	26-OCT-1999;	99US-0161361.	PR	03-JUN-1999;	99US-0137528.
PR	28-OCT-1999;	99US-0161920.	PR	04-JUN-1999;	99US-0137502.
PR	28-OCT-1999;	99US-0161922.	PR	07-JUN-1999;	99US-0137724.
PR	28-OCT-1999;	99US-0161993.	PR	08-JUN-1999;	99US-0138094.
PR	29-OCT-1999;	99US-0162142.	PR	10-JUN-1999;	99US-0138540.
Query Match 68.98; Score 45.5; DB 21; Length 282;					99US-0138847.
Best Local Similarity 77.88; Pred. No. 94;					99US-0139119.
Matches 7; Conservative 0; Mismatches 1; Indels 1; Caps 1;					99US-0139452.
Qy	1 cy-wgcgwyw 8				99US-0139453.
					99US-0139492.
Db	203 cyrwcggyw 211				99US-0139454.
RESULT 15					99US-0139455.
AAAG60113	ID AAG60113 standard; Protein; 285 AA.				99US-0139456.
XX	AC AAG60113;				99US-0139457.
XX	DT 18-OCT-2000 (first entry)				99US-0139458.
XX	DE Arabidopsis thaliana protein fragment SEQ ID NO: 77827.				99US-0139459.
XX	KW Protein identification; signal transduction pathway; metabolic pathway;				99US-0139460.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;				99US-0139461.
KW	termination sequence.				99US-0139462.
XX	XX				99US-0139463.
OS	Arabidopsis thaliana.				99US-0139750.
PN	EPI033405-A2.				99US-0139763.
XX	XX				99US-0139817.
XX	PD 06-SEP-2000.				99US-0139899.
PF	25-FEB-2000; 2000EP-0301439.				99US-0140353.
XX	99US-0121825.				99US-0140354.
PR	05-MAR-1999; 99US-0123180.				99US-0140695.
PR	09-MAR-1999; 99US-0123548.				99US-0140823.

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 68.9%; Score 45.5; DB 21; Length 285;
Best Local Similarity 77.8%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 cy-wgcgyw 8
| | | | |
Db 206 cywgcgyw 214

Search completed: August 12, 2002, 17:21:42
Job time: 199 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:22:10 ; Search time 21.42 Seconds
(without alignments)
9.123 Million cell updates/sec

Title: 10-071247-2

Perfect score: 66

Sequence: 1 cywgcgyw 8

Scoring table: BLOSUM62

Gapex 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2.6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2.6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2.6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2.6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2.6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2.6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	63.6	533	1	US-08-445-586-10
2	42	63.6	533	1	US-08-484-493-13
3	42	63.6	533	1	US-08-484-493-13
4	42	63.6	533	2	US-08-345-212-13
5	42	63.6	533	4	US-09-249-003-13
6	40	60.6	523	1	US-08-100-247-2
7	40	60.6	523	1	US-08-483-146A-2
8	40	60.6	523	1	US-08-232-513A-3
9	40	60.6	523	1	US-08-484-594A-2
10	39	59.1	30	2	US-08-753-829A-9
11	38.5	58.3	469	2	US-08-484-126-1
12	37	56.1	440	4	US-08-759-628-4
13	36	54.5	44	2	US-08-761-248B-12
14	35.5	53.8	28	4	US-08-895-474-13
15	35.5	53.8	31	4	US-08-602-999A-57
16	35.5	53.8	344	3	US-08-278-865-57
17	35.5	53.8	344	3	US-09-110-116-4
18	35.5	53.8	362	1	US-08-415-751-6
19	35	53.0	78	4	US-08-905-223-452
20	35	53.0	430	1	US-08-601-435-2
21	35	53.0	430	2	US-08-931-047-2
22	35	53.0	430	2	US-08-783-202-2
23	35	53.0	552	4	US-08-796-899-28
24	35	53.0	741	3	US-08-943-956A-2
25	35	53.0	2713	5	PCT-US96-01735-1
26	34	51.5	13	4	US-09-258-754-112
27	34	51.5	13	4	US-09-258-754-191

28 34 51.5 13 4 US-09-042-107-112 Sequence 112, App
29 34 51.5 13 4 US-09-042-107-191 Sequence 191, App
30 34 51.5 48 2 US-08-761-248B-7 Sequence 7, Appli
31 34 51.5 212 2 US-08-761-248B-2 Sequence 2, Appli
32 34 51.5 212 2 US-08-761-248B-6 Sequence 6, Appli
33 34 51.5 220 2 US-08-761-248B-4 Sequence 4, Appli
34 34 51.5 292 3 US-08-851-190-3 Sequence 3, Appli
35 34 51.5 333 1 US-08-118-270-12 Sequence 12, Appl
36 34 51.5 333 5 PCT-US93-08528-12 Sequence 12, Appl
37 34 51.5 376 1 US-08-608-241-2 Sequence 2, Appli
38 34 51.5 376 2 US-08-922-182-2 Sequence 2, Appli
39 34 51.5 376 2 US-08-919-953-2 Sequence 2, Appli
40 34 51.5 376 4 US-09-192-983-2 Sequence 2, Appli
41 34 51.5 491 2 US-08-942-819-2 Sequence 2, Appli
42 34 51.5 615 4 US-08-989-299-11 Sequence 11, Appl
43 34 51.5 726 4 US-09-302-812-10 Sequence 10, Appl
44 34 51.5 726 4 US-09-511-477-10 Sequence 10, Appl
45 33 50.0 11 1 US-08-465-391A-339 Sequence 339, App

ALIGNMENTS

RESULT 1
US-08-445-586-10
; Sequence 10, Application US/08445586
; Patent No. 5627050
; GENERAL INFORMATION:
; APPLICANT: Takeshita, Sunao
; APPLICANT: Ito, Toshimi
; APPLICANT: Otawara-Hamamoto, Yoko
; APPLICANT: Amann, Egon
; TITLE OF INVENTION: Bone-Related Sulfatase-Like Protein and
; TITLE OF INVENTION: Process for its Production
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,586
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/111,887
; FILING DATE: 26-AUG-1993
; APPLICATION NUMBER: JP 230030/92
; FILING DATE: 28-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 324034/92
; FILING DATE: 03-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Forman, David S.
; REGISTRATION NUMBER: 33,694
; REFERENCE/DOCKET NUMBER: 02481.1322-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 533 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

US-08-445-586-10

Query Match 63.6%; Score 42; DB 1; Length 533;
Best Local Similarity 85.7%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywgcgyw 8
| | | | |
Db 444 YPGCGYW 450

RESULT 2

US-08-484-493-13
; Sequence 13, Application US/08484493
; Patent No. 5728381
; GENERAL INFORMATION:
; APPLICANT: Wilson, Peter J
; APPLICANT: Morris, Charles P
; APPLICANT: Anson, Donald S
; APPLICANT: Occhiodoro, Teresa
; APPLICANT: Bielicki, Julie
; APPLICANT: Clements, Peter R
; APPLICANT: Hopwood, John J
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; TITLE OF INVENTION: IDURONATE 2-SULFATASE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,493
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 991,973
FILING DATE: 17-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 84162
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 533 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-493-13

Query Match 63.6%; Score 42; DB 1; Length 533;
Best Local Similarity 85.7%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywgcgyw 8
| | | | |
Db 444 YPGCGYW 450

RESULT 3

US-08-484-494-13
; Sequence 13, Application US/08484494
; Patent No. 5798239
; GENERAL INFORMATION:
; APPLICANT: Wilson, Peter J
; APPLICANT: Morris, Charles P
; APPLICANT: Anson, Donald S
; APPLICANT: Occhiodoro, Teresa
; APPLICANT: Bielicki, Julie
; APPLICANT: Clements, Peter R
; APPLICANT: Hopwood, John J
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; TITLE OF INVENTION: IDURONATE 2-SULFATASE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,494
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 991,973
FILING DATE: 17-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 84162
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 533 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-494-13

Query Match 63.6%; Score 42; DB 1; Length 533;
Best Local Similarity 85.7%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywgcgyw 8
| | | | |
Db 444 YPGCGYW 450

RESULT 4

US-08-345-212-13
; Sequence 13, Application US/08345212
; Patent No. 5932211
; GENERAL INFORMATION:
; APPLICANT: Wilson, Peter J
; APPLICANT: Morris, Charles P
; APPLICANT: Anson, Donald S
; APPLICANT: Occhiodoro, Teresa
; APPLICANT: Bielicki, Julie
; APPLICANT: Clements, Peter R
; APPLICANT: Hopwood, John J

1
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; TITLE OF INVENTION: IDURONATE 2-SULFATASE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/345,212
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 991,973
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8416Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 533 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-345-212-13

Query Match 63.6%; Score 42; DB 2; Length 533;
Best Local Similarity 85.7%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywgcgyw 8
| | | | |
Db 444 YPGCGYW 450

RESULT 5
US-09-249-003-13
; Sequence 13, Application US/09249003
; Patent No. 6153188
; GENERAL INFORMATION:
; APPLICANT: Willson, Peter J
; APPLICANT: Morris, Charles P
; APPLICANT: Anson, Donald S
; APPLICANT: Occhiodoro, Teresa
; APPLICANT: Bielicki, Julie
; APPLICANT: Clements, Peter R
; APPLICANT: Hopwood, John J
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; TITLE OF INVENTION: IDURONATE 2-SULFATASE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,003
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/991,973
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8416Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 533 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-249-003-13

Query Match 63.6%; Score 42; DB 4; Length 533;
Best Local Similarity 85.7%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywgcgyw 8
| | | | |
Db 444 YPGCGYW 450

RESULT 6
US-08-100-247-2
; Sequence 2, Application US/08100247
; Patent No. 5571787
; GENERAL INFORMATION:
; APPLICANT: O'BRIEN, JOHN S.
; APPLICANT: KISHIMOTO, YASUO
; TITLE OF INVENTION: PROSAPOSIN AS A NEUTROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
; CITY: NEWPORT BEACH
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/100,247
; FILING DATE: 19930730
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Istaelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: O'BRIEN.002A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
IMMEDIATE SOURCE:
CLONE: PROSAPOSIN
US-08-100-247-2

Query Match 60.6%; Score 40; DB 1; Length 523;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 cywgcgyw 8
| | | | |
DB 494 CIWGPSYW 501

RESULT 7

US-08-483-146A-2
; Sequence 2, Application US/08483146A
; Patent No. 5696080
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; APPLICANT: Kishimoto, Yasuo
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED
; TITLE OF INVENTION: THEREFROM
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Blvd. 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,146A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: MYELOS.002DV1
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: ~ 523 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-483-146A-2

Query Match 60.6%; Score 40; DB 1; Length 523;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 cywgcgyw 8
| | | | |
DB 494 CIWGPSYW 501

RESULT 8

US-08-232-513A-3
; Sequence 3, Application US/08232513A
; Patent No. 5700909
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides
; TITLE OF INVENTION: as Therapeutic Agents
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,513A
; FILING DATE: 21-APR-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/100,247
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 1643
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..523
; OTHER INFORMATION: /label= Hum_prosaposin
US-08-232-513A-3

Query Match 60.6%; Score 40; DB 1; Length 523;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 cywgcgyw 8
| | | | |
DB 494 CIWGPSYW 501

RESULT 9

US-08-484-594A-2
; Sequence 2, Application US/08484594A
; Patent No. 5714459
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; APPLICANT: Kishimoto, Yasuo
; TITLE OF INVENTION: USE OF PROSAPOSIN AND NEUROTROPHIC PEPTIDES
; TITLE OF INVENTION: DERIVED THEREFROM
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,594A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/100,247
FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: MYELOS.002DV2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-484-594A-2

Query Match 60.6%; Score 40; DB 1; Length 523;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 cywgcgyw 8
| | | | |
Db 494 CINGPSYW 501

RESULT 10
US-08-753-829A-9
Sequence 9, Application US/08753829A
Patent No. 5869250
GENERAL INFORMATION:
APPLICANT: Julian, Rudolph L.
APPLICANT: Cheng, Xiaojun
TITLE OF INVENTION: METHOD FOR THE IDENTIFICATION OF
NUMBER OF INVENTION: PEPTIDES THAT RECOGNIZE SPECIFIC DNA SEQUENCES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell, Seltzer, Park & Gibson
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 5869250th Carolina
COUNTRY: United States
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/753,829A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-134
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-520-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-753-829A-9

Query Match 59.1%; Score 39; DB 2; Length 30;
Best Local Similarity 62.5%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 cywgcgyw 8
| | | | |
Db 11 CWDGGYW 18

RESULT 11
US-08-484-126-1
Sequence 1, Application US/08484126
Patent No. 5985655
GENERAL INFORMATION:
APPLICANT: Anderson, W. French
APPLICANT: Baltrucki, Leon F.
APPLICANT: Mason, James M.
TITLE OF INVENTION: Targetable Vector Particles
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
ADDRESSEE: Cecchi, Stewart & Olstein
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: New Jersey
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,126
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,347
FILING DATE: 20-OCT-1994
APPLICATION NUMBER: 08/973,307
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lillie, Raymond J.
REGISTRATION NUMBER: 31,778
REFERENCE/DOCKET NUMBER: 271010-281
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Ecotropic gp70 Protein

US-08-484-126-1

Query Match 58.3%; Score 38.5; DB 2; Length 469;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

Qy 2 ywgc-----gyw 8
|:|:|:|:|
Db 174 YWGCETGRAYW 185

RESULT 12

US-08-759-628-4
; Sequence 4, Application US/08759628
; Patent No. 6225446
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W.
; APPLICANT: Rock, Fernando L.
; APPLICANT: Bazan, J. Fernando
; APPLICANT: Kastelein, Robert A.
; TITLE OF INVENTION: MUTATIONAL VARIANTS OF MAMMLIAN PROTEINS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/759,628
; FILING DATE: 05-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/008,574
; FILING DATE: 06-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0552Q
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-852-9196
; TELEFAX: 415-496-1200
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: proteln
US-08-759-628-4

Query Match 56.1%; Score 37; DB 4; Length 440;
Best Local Similarity 71.4%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ywgcgyw 8
|:|:|:|:|
Db 50 YYGCGYW 56

RESULT 13

US-08-761-248B-12
; Sequence 12, Application US/08761248B
; Patent No. 5958735
; GENERAL INFORMATION:

; APPLICANT: ROWLEY,, DAVID R.
; TITLE OF INVENTION: UROGENITAL SINUS DERIVED GROWTH
; FACTOR NUCLEOTIDE AND AMINO ACID SEQUENCES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jenkins & Gilchrist
; STREET: 1100 Louisiana, Suite 1800
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/761,248B
; FILING DATE: 06-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/008,348
; FILING DATE: 07-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Turley, Charles P.
; REGISTRATION NUMBER: 35,723
; REFERENCE/DOCKET NUMBER: 34012.6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713)9513310
; TELEFAX: (713)9513314
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-761-248B-12

Query Match 54.5%; Score 36; DB 2; Length 44;
Best Local Similarity 71.4%; Pred. No. 60;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 cywgcyg 7
|:|:|:|:|
Db 27 CPDGCY 33

RESULT 14

US-08-895-474-13
; Sequence 13, Application US/08895474
; Patent No. 6136957
; GENERAL INFORMATION:
; APPLICANT: Nicola, Nicos A.
; APPLICANT: Gough, Nicholas M.
; APPLICANT: Gearing, David P.
; APPLICANT: Metcalf, Donald
; APPLICANT: King, Julie Ann
; TITLE OF INVENTION: Improvements in Granulocyte-Macrophage
; Colony-Stimulating Factor Receptor and Derivatives Thereof
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Ave., NW, Ste. 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/895,474
FILING DATE: 16-JUL-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Fox, Samuel L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 1256.00300003
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-895-474-13

Query Match 53.8%; Score 35.5; DB 4; Length 28;
Best Local Similarity 28.6%; Pred. No. 48;
Matches 6; Conservative 1; Mismatches 1; Indels 13; Gaps 1;
QY 1 CYWGC-----GYW 8
|:|:|
Db 4 CWWCCFFKVVQTRCKPDHGIW 24

RESULT 15
US-08-602-999A-57
Sequence 57, Application US/08602999A
Patent No. 6184205
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: KAY, Brian K.
APPLICANT: THORN, Judith M.
APPLICANT: OULLIAM, Lawrence A.
APPLICANT: DER, Channing J.
APPLICANT: FOWLKES, Dana M.
APPLICANT: RIDER, James E.
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
ISOLATING AND USING SAME
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,999A
FILING DATE: 16-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-202
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:

LENGTH: 31 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-602-999A-57

Query Match 53.8%; Score 35.5; DB 4; Length 31;
Best Local Similarity 54.5%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
QY 1 CY-----WGC 6
|:|:|
Db 3 CYREKDTWGC 13

Search completed: August 12, 2002, 17:22:11
Job time: 198 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:13:22 ; Search time 24.99 seconds
(without alignments)
30.761 Million cell updates/sec

Title: 10-071247-1
Perfect score: 59
Sequence: 1 gywgkgyw 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 603

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	42.4	5	2 JH0253	gut pentapeptide -
2	22	37.3	7	2 E33932	Ig mu chain D regi
3	20	33.9	6	2 PT0629	T-cell receptor be
4	20	33.9	7	2 PT0728	T-cell receptor be
5	19	32.2	6	2 A41946	T-cell receptor ga
6	17	28.8	4	2 B53284	T-cell receptor be
7	17	28.8	6	2 A61068	locustakinin - mlg
8	17	28.8	6	2 PT0637	T-cell receptor be
9	17	28.8	7	2 S57274	triacylglycerol li
10	17	28.8	7	2 S33567	tubulin beta-3 cha
11	17	28.8	7	2 PT0628	T-cell receptor be
12	17	28.8	7	2 PT0642	T-cell receptor be
13	17	28.8	7	2 PT0722	T-cell receptor be
14	17	28.8	7	2 PT0586	T-cell receptor be
15	17	28.8	7	2 A38671	peptidylglycine mo
16	17	28.8	7	2 B48394	major fat-globule
17	17	28.8	7	2 P00029	pev-kinin I - pena
18	17	28.8	7	4 I55382	hypothetical pepti
19	17	28.8	8	2 A31570	angiotensin-conver
20	17	28.8	8	2 JS0315	leucokinin V - Mad
21	17	28.8	8	2 JS0316	leucokinin VI - Ma
22	17	28.8	8	2 JS0317	leucokinin VII - M
23	17	28.8	8	2 JS0318	leucokinin VIII -
24	17	28.8	8	2 PT0724	T-cell receptor be
25	15	25.4	6	2 PT0532	T-cell receptor be
26	15	25.4	8	2 S63493	dissimilatory sulf
27	14	23.7	4	2 S09478	globulin IV alpha
28	14	23.7	5	2 PT0278	Ig heavy chain CRD
29	14	23.7	5	2 PT0608	T-cell receptor be

30	14	23.7	6	2 I51434	H4 histone - Afric
31	14	23.7	6	2 B35640	cerebellar degener
32	14	23.7	7	2 S21230	dermorphin (irp-4,
33	14	23.7	8	2 PT0279	Ig heavy chain CRD
34	13	22.0	6	2 A61411	ameletin - rat
35	13	22.0	6	2 PT0514	T-cell receptor be
36	13	22.0	7	1 A61324	dermorphin - Rohde
37	13	22.0	7	2 A60139	fatty-acid synthas
38	13	22.0	7	2 S71870	glutathione transf
39	13	22.0	7	2 S36662	dermorphin (Lys-7)
40	13	22.0	7	2 PT0543	T-cell receptor be
41	13	22.0	7	2 PT0688	T-cell receptor be
42	13	22.0	7	2 PT0579	T-cell receptor be
43	13	22.0	7	2 PT0671	T-cell receptor be
44	13	22.0	8	2 A41117	acetylcholinestera
45	13	22.0	8	2 PH1618	Ig H chain V-D-J r

ALIGNMENTS

RESULT 1

JH0253

gut pentapeptide - Japanese eel

C:Species: Anquilla japonica (Japanese eel)

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Apr-1995

C:Accession: JH0253

R:Uesaka, T.; Ikeda, T.; Kubota, I.; Muneoka, Y.; Ando, M.

Biochem. Biophys. Res. Commun. 180, 828-832, 1991

A:Title: Structure and function of a pentapeptide isolated from the gut of the eel.

A:Reference number: JH0253; MUID:92062113

A:Accession: JH0253

A:Molecule type: protein

A:Residues: 1-5 <UES>

A:Experimental source: gut

C:Comment: This peptide increased basal tone of the circular muscle of the esophagoga

, and of the circular muscle of the gastro-intestinal junction.

Query Match 42.4%; Score 25; DB 2; Length 5;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 gywgk 5

Db 1 GFWNK 5

RESULT 2

E33932

Ig mu chain D region (E7) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 16-Aug-1996

C:Accession: E33932

R:Baccala, R.; Vo Quang, T.; Gilbert, M.; Ternynck, T.; Avrameas, S.

Proc. Natl. Acad. Sci. U.S.A. 86, 4624-4628, 1989

A:Title: Two murine natural polyclonal antibodies are encoded by nonmutated ger

A:Reference number: A33932; MUID:89282823

A:Accession: E33932

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-7 <BAC>

A:Cross-references: GB:M27106

C:Keywords: immunoglobulin

Query Match 37.3%; Score 22; DB 2; Length 7;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 ywgkg 6

Db 3 FYGKG 7

RESULT 3			A:Molecule type: DNA		
PT0629			A:Residues: 1-6 <WHE>		
T-cell receptor beta chain V-D-J region (100-2AH) - mouse (fragment)			C:Keywords: T-cell receptor		
C:Species: Mus musculus (house mouse)			Query Match		
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997			Best Local Similarity 32.2%; Score 19; DB 2; Length 6;		
C:Accession: PT0629; PT0528			Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;		
R:Feeney, A.J.			QY 2 ywgkgy 7		
J. Exp. Med. 174, 115-124, 1991					
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.			Db 1 YRGS GF 6		
A:Reference number: PT0509; MUID:91277601			RESULT 6		
A:Accession: PT0629			B53284		
A:Status: translation not shown			T-cell receptor beta 2 chain D region, Dbeta2 - rabbit		
A:Molecule type: mRNA			C:Species: Oryctolagus cuniculus (domestic rabbit)		
A:Residues: 1-6 <FEE>			C:Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999		
A:Experimental source: newborn thymus, strain BALB/c, clone 100-2AH			C:Accession: B53284		
A:Accession: PT0528			R:Harindranath, N.; Alexander, C.B.; Mage, R.G.		
A:Status: translation not shown			Mol. Immunol. 28, 881-888, 1991		
A:Molecule type: mRNA			A:Title: Evolutionarily conserved organization and sequences of germline diversity an		
A:Residues: 1-6 <FEE>			A:Reference number: A53284; MUID:91342695		
A:Experimental source: adult thymus, strain BALB/c, clone 100-4AB			A:Accession: B53284		
C:Keywords: T-cell receptor			A:Status: preliminary		
Query Match			A:Molecule type: DNA		
Best Local Similarity 33.9%; Score 20; DB 2; Length 6;			A:Residues: 1-4 <HAR>		
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			A:Cross-references: GR:S60737; NID:g233916; PIDN:AA819518.1; PID:g233918		
Qy 1 gywg 4			A:Note: sequence extracted from NCBI backbone (NCBIN:60737, NCBIP:60738)		
			C:Keywords: T-cell receptor		
Db 3 GDWG 6			Query Match		
RESULT 4			Best Local Similarity 28.8%; Score 17; DB 2; Length 4;		
PT0728			Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
T-cell receptor beta chain V-D-J region (161-2H) - mouse (fragment)			QY 3 wg 4		
C:Species: Mus musculus (house mouse)					
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997			Db 2 WG 3		
C:Accession: PT0728			RESULT 7		
R:Feeney, A.J.			A61068		
J. Exp. Med. 174, 115-124, 1991			locustakinin - migratory locust		
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.			C:Species: Locusta migratoria (migratory locust)		
A:Reference number: PT0509; MUID:91277601			C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 08-Dec-1995		
A:Accession: PT0728			C:Accession: A61068		
A:Status: translation not shown			R:Schoofs, L.; Holman, G.M.; Proost, P.; Van Damme, J.; Hayes, T.K.; De Loof, A.		
A:Molecule type: DNA			Regul. Pept. 37, 49-57, 1992		
A:Residues: 1-7 <FEE>			A:Title: Locustakinin, a novel myotropic peptide from Locusta migratoria, isolation,		
A:Experimental source: newborn thymus, strain BALB/c			A:Reference number: A61068; MUID:92262851		
C:Keywords: T-cell receptor			A:Accession: A61068		
Query Match			A:Molecule type: protein		
Best Local Similarity 33.9%; Score 20; DB 2; Length 7;			A:Residues: 1-6 <SCH>		
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			C:Keywords: amidated carboxyl end; cephalomyotropic peptide; neuropeptide		
Qy 1 gywg 4			F;6/Modified site: amidated carboxyl end (Gly) #status experimental		
			Query Match		
Db 3 GDWG 6			Best Local Similarity 28.8%; Score 17; DB 2; Length 6;		
RESULT 5			Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
A41946			QY 3 wg 4		
T-cell receptor gamma chain (1t.22) - mouse (fragment)					
C:Species: Mus musculus (house mouse)			Db 5 WG 6		
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999			RESULT 8		
C:Accession: A41946			PT0637		
R:Whetsell, W.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.			T-cell receptor beta chain V-D-J region (111-1K) - mouse (fragment)		
Mol. Cell. Biol. 11, 5902-5909, 1991			C:Species: Mus musculus (house mouse)		
A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge			C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997		
A:Reference number: A41946; MUID:92049316					
A:Accession: A41946					
A:Status: preliminary; not compared with conceptual translation					

C;Accession: PT0637
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0637
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 28.8%; Score 17; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
||
Db 5 WG 6

RESULT 9

S5727A
triacylglycerol lipase (BC 3.1.1.3) - Psychrobacter immobilis (fragment)
C;Species: Psychrobacter immobilis
C;Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 03-Nov-1995
A;Accession: S5727A
R;Arpigny, J.L.; Feller, G.; Garday, C.
Biochim. Biophys. Acta 1263, 103, 1995
A;Title: Corrigendum to "Cloning, sequence and structural features of a lipase from the
A;Reference number: S57274; MUID:95359197
A;Accession: S5727A
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-7 <ARP>
A;Cross-references: EMBL:X67712
C;Keywords: carboxylic ester hydrolase

Query Match 28.8%; Score 17; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
||
Db 1 WG 2

RESULT 10

S33567
tubulin beta-3 chain - fruit fly (Drosophila melanogaster) (fragment)
C;Species: Drosophila melanogaster
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 21-Jul-2000
A;Accession: S33567
R;Hinz, U.; Wolk, A.; Renkawitz-Pohl, R.
Development 116, 543-554, 1992
A;Title: Ultrathiorax is a regulator of beta-3 tubulin expression in the Drosophila vis
A;Reference number: S33567; MUID:93170162
A;Accession: S33567
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-7 <HIN>
A;Cross-references: EMBL:X68393; NID:g7669; PIDN:CAA48459.1; PID:g4377448
C;Genetics:
A;Gene: FlyBase:tbgr
A;Cross-references: FlyBase:FBgn0003888
A;Introns: 5/3

Query Match 28.8%; Score 17; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 gkgyw 8
| :|
Db 3 GAKFW 7

RESULT 11

PT0628
T-cell receptor beta chain V-D-J region (111-LAG) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0628
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0628
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 28.8%; Score 17; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
||
Db 6 WG 7

RESULT 12

PT0642
T-cell receptor beta chain V-D-J region (111-LH) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0642
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0642
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 28.8%; Score 17; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
||
Db 5 WG 6

RESULT 13

PT0722
T-cell receptor beta chain V-D-J region (135-LG) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0722
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0722
A;Status: translation not shown
A;Molecule type: DNA

A;Residues: 1-7 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 28.8%; Score 17; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
||
Db 6 WG 7 .

RESULT 14

PT0586
T-cell receptor beta chain V-D-J region (141-1CN) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0586; PT0592
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0586
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FEE>
A;Experimental source: day 19 fetal thymus, strain BALB/c (clones 141-1CN and 141-1CD)
C;Keywords: T-cell receptor

Query Match 28.8%; Score 17; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
||
Db 5 WG 6

RESULT 15

A38671
peptidylglycine monooxygenase (EC 1.14.17.3) form 3 - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 23-Jun-1993
C;Accession: A38671
R;Stoffers, D.A.; Ouafik, L.; Eipper, B.A.
J. Biol. Chem. 266, 1701-1707, 1991
A;Title: Characterization of novel mRNAs encoding enzymes involved in peptide alpha-amidation
A;Reference number: A38671; MUID:91107670
A;Accession: A38671
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-7 <STO>
C;Keywords: oxidoreductase

Query Match 28.8%; Score 17; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 gkg 6
|||
Db 4 GKG 6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:15:03 ; Search time 13.13 Seconds
(without alignments)
23.592 Million cell updates/sec

Title: 10-071247-1
Perfect score: 59
Sequence: 1 gywgkgyw 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 148

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	28.8	6	1 LOK1_LOCM1	P41491 locusta mlg
2	17	28.8	8	1 ACL_THUAL	P18691 thunnus alb
3	17	28.8	8	1 LCK1_LEUMA	P21140 leucophaea
4	17	28.8	8	1 LCK2_LEUMA	P21141 leucophaea
5	17	28.8	8	1 LCK3_LEUMA	P21142 leucophaea
6	17	28.8	8	1 LCK4_LEUMA	P21143 leucophaea
7	17	28.8	8	1 LCK5_LEUMA	P19987 leucophaea
8	17	28.8	8	1 LCK6_LEUMA	P19988 leucophaea
9	17	28.8	8	1 LCK7_LEUMA	P19989 leucophaea
10	17	28.8	8	1 LCK8_LEUMA	P19990 leucophaea
11	15	25.4	4	1 OCP3_OCTMI	P58649 octopus min
12	15	25.4	5	1 UF01_MOUSE	P38639 mus musculus
13	15	25.4	8	1 ALL5_CYPDPO	P82156 cydia pomon
14	12	20.3	8	1 CCKN_MACEU	P30369 macropus eu
15	12	20.3	8	1 VGLG_HSV2B	P81780 herpes simp
16	11	18.6	5	1 BPP7_BOTIN	P30425 bothrops in
17	11	18.6	6	1 ET01_LITRU	P82096 litoria rub
18	11	18.6	7	1 FAR5_HIRME	P42564 hiruudo medi
19	11	18.6	7	1 TY51_LITRU	P82065 litoria rub
20	11	18.6	7	1 WNA1_ACHFU	P35919 achatina fu
21	11	18.6	7	1 WNA2_ACHFU	P35920 achatina fu
22	11	18.6	7	1 WNA3_ACHFU	P35921 achatina fu
23	11	18.6	8	1 AKHG_GRYBI	P14086 gryllus bim
24	11	18.6	8	1 AKH_LIBAU	P25418 libellula a
25	11	18.6	8	1 AKH_MELML	P25423 melolontha
26	11	18.6	8	1 AKH_TABAT	P14595 tabanus atr
27	11	18.6	8	1 HTF1_PERAM	P04548 periplaneta
28	11	18.6	8	1 HTF2_PERAM	P04549 periplaneta
29	11	18.6	8	1 HTF_TENMO	P25419 tenebrio mo
30	11	18.6	8	1 PLP_BRANA	P81707 brassica na
31	11	18.6	8	1 RECH_PANBO	P08939 pandanus bo
32	11	18.6	8	1 RSI_ERWCH	P37985 erwania chr
33	11	18.6	8	1 RS7_MYCIT	P33564 mycobacteri

34	11	18.6	8	1 UF06_MOUSE	P38644 mus musculus
35	10	16.9	5	1 PAP2_PARMA	P81864 pardachirus
36	10	16.9	5	1 RE32_LITRU	P82073 litoria rub
37	10	16.9	6	1 ASP2_LACSN	P82655 lactobacill
38	10	16.9	8	1 AL17_CARMA	P81820 carcinus ma
39	10	16.9	8	1 ALL5_CALVO	P41841 calliphora
40	9	15.3	4	1 ACHI_ACHFU	P35904 achatina fu
41	9	15.3	4	1 OCP1_OCTMI	P58648 octopus min
42	9	15.3	7	1 ALL7_CYPDPO	P82158 cydia pomon
43	9	15.3	7	1 MNPI_LBPDE	P42984 leptinotars
44	9	15.3	8	1 AL16_CARMA	P81819 carcinus ma
45	9	15.3	8	1 ALL9_CARMA	P81812 carcinus ma

ALIGNMENTS

RESULT 1				
LOK1_LOCM1				
ID	LOK1_LOCM1	STANDARD;	PRT;	6 AA.
AC	P41491;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	Locustakinin I.			
OS	Locusta migratoria (Migratory locust).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;			
OC	Acridomorpha; Acridoidea; Acrididae; Locusta.			
OX	NCBI_TaxID=7004;			
[1]				
RN	SEQUENCE.			
RP				
RC	TISSUE=Corpora cardiaca;			
RX	MEDLINE=92262851; PubMed=1585017;			
RA	Schoofs L., Holman G.M., Proost P., van Damme J., Hayes T.K.,			
RA	de Loof A.;			
RT	"Locustakinin, a novel myotropic peptide from Locusta migratoria,			
RT	isolation, primary structure and synthesis.";			
RL	Regul. Pept. 37:49-57(1992).			
CC	-1- FUNCTION: MYOTROPIC PEPTIDE. MAY BE IMPORTANT IN THE STIMULATION			
CC	OF ION TRANSPORT AND INHIBITION OF DIURETIC ACTIVITY IN MALPIGHIAN			
CC	TUBULES.			
DR	PIR: A61058; A61068.			
KW	Neuropeptide; Amidation.			
FT	MOD_RES 6 6			
SQ	SEQUENCE 6 AA; 654 MW; 686365A5B9CDB000 CRC64;			

Query Match 28.8%; Score 17; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
||
Db 5 WG 6

RESULT	2
ACI_THUAL	
ID	ACI_THUAL
AC	P1869
DT	01-NO
DT	01-NO
DT	01-NO
DE	Angio
OS	Thunn
OC	Eukar
OC	Actin
OC	Acant
OC	Scomb
OX	NCBI_
RN	[1]
RP	SEQU

```
RC TISSUE=Muscle;
RX MEDLINE=88326322; PubMed=3415688;
RA Kohana Y., Matsumoto S., Oka H., Teramoto T., Okabe M., Mimura T.;
RT "Isolation of angiotensin-converting enzyme inhibitor from tuna
RT muscle.";
RL Biochem. Biophys. Res. Commun. 155:332-337(1988).
DR PIR: A31570; A31570.
SQ SEQUENCE 8 AA; 953 MW; 6AA863733051F1B7 CRC64;

Query Match 28.8%; Score 17; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
DB 6 wg 7

RESULT 3
LCK1_LEUMA STANDARD; PRT; 8 AA.
ID LCK1_LEUMA
AC P21140;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DE Leucokinin I (L-I).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Head;
RA Holman G.N., Cook B.J., Nachman R.J.;
RT "Isolation, primary structure and synthesis of two neuropeptides
RT from Leucophaea maderae: members of a new family of
RT Cephalomyotropins.";
RL Comp. Biochem. Physiol. 84C:205-211(1986).
CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC -!- ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 893 MW; DC6365B449CDC76A CRC64;

Query Match 28.8%; Score 17; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
DB 7 wg 8

RESULT 4
LCK2_LEUMA STANDARD; PRT; 8 AA.
ID LCK2_LEUMA
AC P21141;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DE Leucokinin II (L-II).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Head;
RA Holman G.N., Cook B.J., Nachman R.J.;
RT "Isolation, primary structure and synthesis of two additional neuropeptides
RT from Leucophaea maderae: members of a new family of
RT Cephalomyotropins.";
RL Comp. Biochem. Physiol. 84C:271-276(1986).
CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC -!- ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 910 MW; DC6365B449C866DA CRC64;

Query Match 28.8%; Score 17; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
DB 7 wg 8

RESULT 6
LCK4_LEUMA STANDARD; PRT; 8 AA.
ID LCK4_LEUMA
AC P21143;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DE Leucokinin IV (L-IV).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
```

```
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Isolation, primary structure and synthesis of two neuropeptides
RT from Leucophaea maderae: members of a new family of
RT Cephalomyotropins.";
RL Comp. Biochem. Physiol. 84C:205-211(1986).
CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC -!- ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 852 MW; DC6365A5B9C8676A CRC64;

Query Match 28.8%; Score 17; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
DB 7 wg 8

RESULT 5
LCK3_LEUMA STANDARD; PRT; 8 AA.
ID LCK3_LEUMA
AC P21142;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DE Leucokinin III (L-III).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Head;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Primary structure and synthesis of two additional neuropeptides
RT from Leucophaea maderae: members of a new family of
RT Cephalomyotropins.";
RL Comp. Biochem. Physiol. 84C:271-276(1986).
CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC -!- ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 910 MW; DC6365B449C866DA CRC64;

Query Match 28.8%; Score 17; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
DB 7 wg 8

RESULT 6
LCK4_LEUMA STANDARD; PRT; 8 AA.
ID LCK4_LEUMA
AC P21143;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DE Leucokinin IV (L-IV).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
```

```

RN SEQUENCE, AND SYNTHESIS.
RC TISSUE-Head;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Primary structure and synthesis of two additional neuropeptides
RT from Leucophaea maderae: members of a new family of
RL Cephalomyotropins.";
RL Comp. Biochem. Physiol. 84C:271-276(1986).
CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 906 MW; DC6365B1E9D5BDDA CRC64;

Query Match 28.8%; Score 17; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
DB 7 WG 8

RESULT 7
LCK5_LEUMA
ID LCK5_LEUMA STANDARD; PRT; 8 AA.
AC P19987;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Leucokinin V (L-V).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE.
RC TISSUE-Head;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Isolation, primary structure, and synthesis of leucokinin V and VI:
RT myotropic peptides of Leucophaea maderae.";
RL Comp. Biochem. Physiol. 88C:27-30(1987).
CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.
DR PIR; JS0315;
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 784 MW; 736365A5B9C865B8 CRC64;

Query Match 28.8%; Score 17; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
DB 7 WG 8

RESULT 8
LCK6_LEUMA
ID LCK6_LEUMA STANDARD; PRT; 8 AA.
AC P19988;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Leucokinin VI (L-VI).
OS Leucophaea maderae (Madeira cockroach).
```

```

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE.
RC TISSUE-Head;
RX MEDLINE=87052651; PubMed=2877794;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Isolation, primary structure, and synthesis of leucokinin V and VI:
RT myotropic peptides of Leucophaea maderae.";
RL Comp. Biochem. Physiol. 88C:27-30(1987).
CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
CC -!- SIMILARITY: TO THE OTHER LEUCOKININS, AND TO MANDUCA SEXTA AND
CC HELIOTHIS ZEA ADIPOKINETIC HORMONE.
CC PIR; JS0316; JS0316.
DR Neuropeptide; Amidation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 935 MW; 9D6365B1E9D5A5A6 CRC64;

Query Match 28.8%; Score 17; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
DB 7 WG 8

RESULT 9
LCK7_LEUMA
ID LCK7_LEUMA STANDARD; PRT; 8 AA.
AC P19989;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Leucokinin VII (L-VII).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE.
RC TISSUE-Head;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Isolation, primary structure and synthesis of leucokinin VII and
RT VIII; the final members of this new family of cephalomyotropic
RT peptides isolated from head extracts of Leucophaea maderae.";
RL Comp. Biochem. Physiol. 88C:31-34(1987).
CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.
DR PIR; JS0317; JS0317.
KW Neuropeptide; Amidation.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 866 MW; DC6365A5B9C8D76A CRC64;

Query Match 28.8%; Score 17; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
DB 7 WG 8

RESULT 10
LCK8_LEUMA
```

```

ID LCK8_LEUMA STANDARD; PRT; 8 AA.
AC P19990;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE Leucokinin VIII (L-VIII).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE.
RC TISSUE=Head;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Isolation, primary structure and synthesis of leucokinin VII and
RT VIII: the final members of this new family of cephalomyotropic
RT peptides isolated from head extracts of Leucophaea maderae.";
RL Comp. Biochem. Physiol. 88C:31-34(1987).
CC -1- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
CC -1- SIMILARITY: TO THE OTHER LEUCOKININS.
DR PIR; JS0318; JS0318.
KW Neuropeptide; Amidation.
FT MOD_RES 8 8
SQ SEQUENCE 8 AA; 902 MW; 736365AB59CAADD8 CRC64;

Query Match 28.8%; Score 17; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
DB 7 wg 8

RESULT 11
OCP3_OCTMI STANDARD; PRT; 4 AA.
AC P58649;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Cardiac active peptides Ocp-3/Ocp-4.
OS Octopus minor (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
OC Incirrata; Octopodidae; Octopus.
OX NCBI_TaxID=89766;
RN [1]
RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
RC TISSUE=Brain;
RX PubMed=10876044;
RA Iwakoshi E., Hisada M., Minakata H.;
RT "Cardioactive peptides isolated from the brain of a Japanese octopus,
RT Octopus minor.";
RL Peptides 21:623-630(2000).
CC -1- FUNCTION: Cardioactive; has both positive chronotropic and
CC inotropic effects on the heart. Ocp-4 is a 1000 time less
CC active than Ocp-3.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- PTM: Ocp-4 has D-Ser instead of L-Ser.
CC -1- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.
KW Hormone; D-amino acid.
FT MOD_RES 2 2
SQ SEQUENCE 4 AA; 463 MW; 6AB365B810000000 CRC64;

Query Match 25.4%; Score 15; DB 1; Length 4;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

QY 1 gvw 3
DB 7 gvw 8

```

```

DB 1 GSW 3

RESULT 12
UF01_MOUSE STANDARD; PRT; 5 AA.
ID UF01_MOUSE
AC P38639;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Unknown protein from 2D-page of fibroblasts (P19) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC TISSUE=Fibroblast; PubMed=7523108;
RX MEDLINE=95009907; Wichter L.L., He C., Selkirk J.K.;
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT "Separation and sequencing of familial and novel murine proteins
RT using preparative two-dimensional gel electrophoresis.";
RL Electrophoresis 15:735-745(1994).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.6, ITS MW IS: 19 kDa.
FT NON_TER 5 5
SQ SEQUENCE 5 AA; 717 MW; 7364087043100000 CRC64;

Query Match 25.4%; Score 15; DB 1; Length 5;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gvw 3
DB 3 GRW 5

RESULT 13
ALL5_CYDPO STANDARD; PRT; 8 AA.
ID ALL5_CYDPO
AC P82156;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydia statin 5.
OS Cydia pomonella (Codling moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricoidea; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82600;
RN [1]
RP SEQUENCE.
RC TISSUE=Larva; PubMed=9392829;
RX MEDLINE=98054539; Maestri J.-L., Scott A.G., Winstanley D.,
RA Duve H., Johnson A.H., Maestri J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8 8
SQ SEQUENCE 8 AA; 898 MW; 922879CABB58640D CRC64;

Query Match 25.4%; Score 15; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 kgy 7
DB 2 RGY 4

```



```
RESULT 14
CCKN_MACEU
ID CCKN_MACEU STANDARD; PRT; 8 AA.
AC P30369;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Cholecystokinin (CCK).
GN CCK
OS Macropus eugenii (Tammam wallaby), and
OS Dasyurus viverrinus (Southeastern quoll).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9315, 9279;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=88234141; PubMed=3375140;
RT Fan Z.W., Eng J., Shaw G., Yalow R.S.;
RT "Cholecystokinin octapeptide purified from brains of Australian
RT marsupials.";
RL Peptides 9:429-431(1988).
CC -!- FUNCTION: THIS PEPTIDE HORMONE INDUCES GALL BLADDER CONTRACTION
CC AND THE RELEASE OF PANCREATIC ENZYMES IN THE GUT. ITS FUNCTION
CC IN THE BRAIN IS NOT CLEAR.
CC -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
DR PIR; A43001; A43001.
DR PIR; P00012; P00012.
DR InterPro; IPR001651; Gastrin.
DR PROSITE; PS00259; GASTRIN; 1.
KW Amidation; Sulfation; Hormone.
FT MOD_RES 2 SULFATION.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 1064 MW; DDCAA68378768B5A CRC64;

Query Match 20.3%; Score 12; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywg 4
Db 2 YMG 4

RESULT 15
VGLG_HSV2B
ID VGLG_HSV2B STANDARD; PRT; 8 AA.
AC P81780;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Glycoprotein G (Fragment).
OS Herpes simplex virus (type 2 / strain B4327UR).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=103921;
RN [1]
RP SEQUENCE.
RA Liljeqvist J.-A., Svennerholm B., Bergstrom T.;
RL Submitted (APR-1999) to the SWISS-PROT data bank.
CC -!- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
CC 2: GH, GG, GC, GD, GI, AND GE.
CC -!- MISCELLANEOUS: GLYCOPROTEIN G IS MUCH LARGER IN HSV-2 THAN IN
CC HSV-1.
KW Glycoprotein.
FT NON_TER 8
SQ SEQUENCE 8 AA; 683 MW; 7B47686772C865B8 CRC64;

Query Match 20.3%; Score 12; DB 1; Length 8;
```

Best Local Similarity 66.7%; Pred. No. 1e+05; 1; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 gkg 6
Db 1 GSG 3
Search completed: August 12, 2002, 17:20:43
Job time: 340 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:14:38 ; Search time 39.45 Seconds
(without alignments)
35.081 Million cell updates/sec

Title: 10-071247-1
Perfect score: 59
Sequence: 1 gywgkgyw 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 segs, 172994929 residues
Total number of hits satisfying chosen parameters: 378

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL19:**

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	28.8	7	10	O49223 glycine max
2	17	28.8	8	5	P82685 periplaneta
3	17	28.8	8	5	P82686 periplaneta
4	17	28.8	8	5	P82687 periplaneta
5	17	28.8	8	5	P82688 periplaneta
6	17	28.8	8	5	P82689 periplaneta
7	17	28.8	8	6	P82929 bos taurus
8	15	25.4	8	4	Q96RN9
9	15	25.4	8	12	O64971 alfalfa mos
10	13	22.0	8	2	Q95822
11	13	22.0	8	4	Q15888
12	13	22.0	8	6	O02831
13	13	22.0	8	7	Q95213
14	12	20.3	7	8	Q95945
15	12	20.3	8	4	Q9UCN4
16	11.5	19.5	8	11	Q62721 rattus norv

17	11	18.6	6	13	P82096
18	11	18.6	7	2	Q47029
19	11	18.6	7	13	P82065
20	11	18.6	8	2	O09258
21	11	18.6	8	2	O85406
22	11	18.6	8	2	O44463
23	11	18.6	8	2	O56429
24	11	18.6	8	2	Q93SP2
25	11	18.6	8	4	Q15890
26	11	18.6	8	5	Q9VRD2
27	11	18.6	8	6	Q9TRY3
28	11	18.6	8	6	Q95M23
29	11	18.6	8	8	Q9TD02
30	11	18.6	8	8	Q35792
31	11	18.6	8	8	Q9T4Y2
32	11	18.6	8	8	Q94VJ4
33	11	18.6	8	8	Q94VF9
34	11	18.6	8	8	Q94VF6
35	11	18.6	8	8	Q94VE4
36	11	18.6	8	8	Q94VC1
37	11	18.6	8	8	Q94VB5
38	11	18.6	8	8	Q94VB2
39	11	18.6	8	8	Q94VA7
40	11	18.6	8	8	Q94V91
41	11	18.6	8	8	Q94V88
42	11	18.6	8	8	Q94V82
43	11	18.6	8	8	Q94PX7
44	11	18.6	8	8	Q94PX6
45	11	18.6	8	8	Q94PX5

ALIGNMENTS

RESULT 1

O49223 ID O49223 PRELIMINARY; PRT; 7 AA.
AC O49223;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HMG-1-LIKE PROTEIN (FRAGMENT).
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ESSEX; TISSUE=ROOT;
RX MEDLINE=91367679; Pubmed=1891369;
RA Laux T., Goldberg R.B.;
RT "A plant DNA binding protein shares highly conserved sequence motifs
with HMG-box proteins.";
RL Nucleic Acids Res. 19:4769-4769(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ESSEX; TISSUE=ROOT;
RA Mahalingam R., Knap H.T.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF047050; AAC03556.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 850 MW; 6AAAAAB378637810 CRC64;

Query Match 28.8%; Score 17; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 wg 4

Db 2 WG 3

RESULT 2
P82685 ID P82685 PRELIMINARY; PRT; 8 AA.
AC P82685;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE KININ-1 (PEA-K-1).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattoidea; Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=CORPORA CARDIACA;
RX MEDLINE=98010462; PubMed=9350979;
RA Predel R., Kellner R., Rapus J., Penzlin H., Gaede G.;
RT "Isolation and structural elucidation of eight kinins from the retrocerebral complex of the American cockroach, Periplaneta americana.";
RL Regul. Pept. 71:199-205(1997).
CC -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC -!- (MYOTROPIC ACTIVITY).
CC -!- SIMILARITY: BELONGS TO THE KININ FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 950 MW; 376365B449D5A774 CRC64;

Query Match 28.8%; Score 17; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
||
Db 7 WG 8

RESULT 3
P82686 ID P82686 PRELIMINARY; PRT; 8 AA.
AC P82686;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE KININ-2 (PEA-K-2).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattoidea; Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=CORPORA CARDIACA;
RX MEDLINE=98010462; PubMed=9350979;
RA Predel R., Kellner R., Rapus J., Penzlin H., Gaede G.;
RT "Isolation and structural elucidation of eight kinins from the retrocerebral complex of the American cockroach, Periplaneta americana.";
RL Regul. Pept. 71:199-205(1997).
CC -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC -!- (MYOTROPIC ACTIVITY).
CC -!- SIMILARITY: BELONGS TO THE KININ FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 856 MW; DC6365A5B9D5EDDA CRC64;

Query Match 28.8%; Score 17; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
||
Db 7 WG 8

RESULT 4
P82687 ID P82687 PRELIMINARY; PRT; 8 AA.
AC P82687;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE KININ-3 (PEA-K-3).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattoidea; Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=CORPORA CARDIACA;
RX MEDLINE=98010462; PubMed=9350979;
RA Predel R., Kellner R., Rapus J., Penzlin H., Gaede G.;
RT "Isolation and structural elucidation of eight kinins from the retrocerebral complex of the American cockroach, Periplaneta americana.";
RL Regul. Pept. 71:199-205(1997).
CC -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC -!- (MYOTROPIC ACTIVITY).
CC -!- SIMILARITY: BELONGS TO THE KININ FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 909 MW; DC6365B449D5A76A CRC64;

Query Match 28.8%; Score 17; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
||
Db 7 WG 8

RESULT 5
P82688 ID P82688 PRELIMINARY; PRT; 8 AA.
AC P82688;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE KININ-4 (PEA-K-4).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattoidea; Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=CORPORA CARDIACA;
RX MEDLINE=98010462; PubMed=9350979;
RA Predel R., Kellner R., Rapus J., Penzlin H., Gaede G.;
RT "Isolation and structural elucidation of eight kinins from the retrocerebral complex of the American cockroach, Periplaneta americana.";
RL Regul. Pept. 71:199-205(1997).
CC -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC -!- (MYOTROPIC ACTIVITY).
CC -!- SIMILARITY: BELONGS TO THE KININ FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 839 MW; 736365A5B9D6DDDD8 CRC64;

```
Query Match      28.8%; Score 17; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
   ||
Db 7 WG 8

RESULT 6
ID P82689 PRELIMINARY; PRT; 8 AA.
AC P82689;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE KININ-5 (PEA-K-5).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattoidea; Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE AND FUNCTION.
RC TISSUE=CORPORA CARDIACA;
RX MEDLINE=98010462; PubMed=9350979;
RA Predel R., Kellner R., Rapus J., Penzlin H., Gaede G.;
RT "Isolation and structural elucidation of eight kinins from the
RT retrocerebral complex of the American cockroach, Periplaneta
RT americana.";
RL Regul. Pept. 71:199-205(1987).
CC -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC -!- (MYOTROPIC ACTIVITY).
CC -!- SIMILARITY: BELONGS TO THE KININ FAMILY.
KW Neuropeptide; Amidation.
FT MOD.RES 8 8
SQ SEQUENCE 8 AA; 865 MW; C76365B449CDC775 CRC64;

Query Match      28.8%; Score 17; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
   ||
Db 7 WG 8

RESULT 7
ID P82929 PRELIMINARY; PRT; 8 AA.
AC P82929;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE MITOCHONDRIAL 28S RIBOSOMAL PROTEIN S34 (MRP-S34) (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=LIVER;
RA Koc E.C., Burkhardt W., Blackburn K., Moseley A., Spremulli L.L.;
RT "Small subunit of the mammalian mitochondrial ribosome. Identification
RT of the full complement ribosomal proteins present.";
RL Submitted (DEC-2000) to the SWISS-PROT data bank.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
KW Ribosomal protein; Mitochondrion.
FT NON_TER 1 1
```

```
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 935 MW; 9639D1A72058637D CRC64;

Query Match      28.8%; Score 17; DB 6; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
   ||
Db 2 WG 3

RESULT 8
ID Q96RN9 PRELIMINARY; PRT; 8 AA.
AC Q96RN9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TRUNCATED PROTOPORPHYRINOGEN OXIDASE (FRAGMENT).
GN PPOX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21249095; PubMed=11350188;
RA Corrigan A.V., Hift R.J., Davids L.M., Hancock V., Meissner D.,
RA Kirsch R.E., Meissner P.N.;
RT "Identification of the First Variegate Porphyria Mutation in an
RT Indigenous Black South African and Further Evidence for Heterogeneity
RT in Variegate Porphyria.";
RL Mol. Genet. Metab. 73:91-96(2001).
DR EMBL; AF321822; AAK69608.1; -.
FT NON_TER 1 1
SQ SEQUENCE 8 AA; 862 MW; 893365B878772682 CRC64;

Query Match      25.4%; Score 15; DB 4; Length 8;
Best Local Similarity 66.7%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gyw 3
   ||
Db 5 GSW 7

RESULT 9
ID Q64971 PRELIMINARY; PRT; 8 AA.
AC Q64971;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE ORF (FRAGMENT).
OS Alfalfa mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Alfamovirus.
OX NCBI_TaxID=12321;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81124289; PubMed=6927843;
RA Koper-Zwarthoff E.C., Brederode F.T.M., Veeneman G., van Boom J.H.,
RA Bol J.F.;
RT "Nucleotide sequences at the 5'-termini of the alfalfa mosaic virus
RT RNAs and the intercistronic junction in RNA 3.";
RL Nucleic Acids Res. 8:5635-5647(1980).
DR EMBL; V00047; CAA23416.1; -.
FT NON_TER 1 1
SQ SEQUENCE 8 AA; 917 MW; 69D40B0775A365B8 CRC64;
```

```
Query Match      25.4%; Score 15; DB 12; Length 8;
Best Local Similarity 66.7%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gyw 3
   | |
Db 1 GSW 3

RESULT 10
Q9R5R2 PRELIMINARY; PRT; 8 AA.
AC Q9R5R2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 31,000 DA PRODUCT OF ORFB.
OS Shigella dysenteriae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Shigella.
OX NCBI_TaxID=622;
RN [1]
RP SEQUENCE.
RX MEDLINE=92085268; PubMed=1660923;
RA Polard P., Preere M.F., Chandler M., Fayet O.;
RT "Programmed translational frameshifting and initiation at an AUU codon
in gene expression of bacterial insertion sequence IS911.";
RL J. Mol. Biol. 222:465-477(1991).
SQ SEQUENCE 8 AA; 902 MW; FE2DCAFB586AE336 CRC64;

Query Match      22.0%; Score 13; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gy 2
   | |
Db 4 GY 5

RESULT 11
Q15888 PRELIMINARY; PRT; 8 AA.
AC Q15888;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE (CLONE XP15H8A) (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinnault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L32069; AAA73878.1; -.
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 1068 MW; 0315A37EAB5B0763 CRC64;

Query Match      22.0%; Score 13; DB 4; Length 8;
Best Local Similarity 33.3%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 wgt 5

Query Match      25.4%; Score 15; DB 12; Length 8;
Best Local Similarity 66.7%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gyw 3
   | |
Db 1 GSW 3

RESULT 10
Q9R5R2 PRELIMINARY; PRT; 8 AA.
AC Q9R5R2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 31,000 DA PRODUCT OF ORFB.
OS Shigella dysenteriae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Shigella.
OX NCBI_TaxID=622;
RN [1]
RP SEQUENCE.
RX MEDLINE=92085268; PubMed=1660923;
RA Polard P., Preere M.F., Chandler M., Fayet O.;
RT "Programmed translational frameshifting and initiation at an AUU codon
in gene expression of bacterial insertion sequence IS911.";
RL J. Mol. Biol. 222:465-477(1991).
SQ SEQUENCE 8 AA; 902 MW; FE2DCAFB586AE336 CRC64;

Query Match      22.0%; Score 13; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gy 2
   | |
Db 4 GY 5

RESULT 11
Q15888 PRELIMINARY; PRT; 8 AA.
AC Q15888;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE (CLONE XP15H8A) (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinnault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L32069; AAA73878.1; -.
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 1068 MW; 0315A37EAB5B0763 CRC64;

Query Match      22.0%; Score 13; DB 4; Length 8;
Best Local Similarity 33.3%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 wgt 5

Query Match      22.0%; Score 13; DB 6; Length 8;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 yw 3
   | |
Db 1 HW 2

RESULT 13
Q95213 PRELIMINARY; PRT; 8 AA.
AC Q95213;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE GERMLINE DH (DF) GENE (FRAGMENT).
GN DF.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F-I/RGM;
RA Mage R.G., Chen H.-T., Alexander C.B., Chen F.F.;
RT "Rabbit DQ52 and DH Gene Rearrangements in Early B-cell Development.";
RL Mol. Immunol. 0:0-0(1996).
DR EMBL; U62585; AAB18735.1; -.
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 845 MW; 5CA861B5AB58677B CRC64;

Query Match      22.0%; Score 13; DB 7; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gy 2
   | |
Db 3 GY 4
```

RESULT 14
Q95945 PRELIMINARY; PRT; 7 AA.
AC Q95945;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE INSIDE INTRON 5 (FRAGMENT).
OS Saccharomyces cerevisiae (Baker's yeast).
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D273-10B;
RX MEDLINE=81069885; PubMed=6254986;
RA Bonitz S.G., Coruzzi G., Thalenfeld B., Tzagoloff A., Macino G.;
RT "Assembly of the mitochondrial membrane system: Structure and
RT nucleotide sequence of the gene coding for subunit 1 of yeast
RT cytochrome oxidase.";
RL J. Biol. Chem. 255:11927-11941(1980).
DR EMBL: V00694; CAA24066.1; -;
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 7 AA; 859 MW; 75B7232362CDC460 CRC64;

Query Match 20.3%; Score 12; DB 8; Length 7;
Best Local Similarity 25.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 5 Kgyw 8
: |
Db 1 RAVW 4

RESULT 15
Q9UCN4 PRELIMINARY; PRT; 8 AA.
AC Q9UCN4;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE CELL-SURFACE HEPARIN/HEPARANSULFATE-BINDING PROTEIN PEPTIDE 3
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=92291065; PubMed=1601862;
RA Raboudi N., Julian J., Rohde L.H., Carson D.D.;
RT "Identification of cell-surface heparin/heparan sulfate-binding
RT proteins of a human uterine epithelial cell line (RL95).";
RL J. Biol. Chem. 267:11930-11939(1992).
SQ SEQUENCE 8 AA; 689 MW; 80B8733DD33DD87D CRC64;
Query Match 20.3%; Score 12; DB 4; Length 8;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 gywgkg 6
: | |
Db 2 GAKAKG 7

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:10:27 ; Search time 48.76 Seconds
(without alignments)
18,224 Million cell updates/sec

Title: 10-071247-1
Perfect score: 59
Sequence: 1 gywgkgyw 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 69368

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*

- 1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	8	20	AA1980 (F-18)
2	59	100.0	8	20	AA1981 (F-18)
3	51	86.4	8	21	AA1982 (F-18)
4	45	76.3	8	20	AA1983 (F-18)
5	45	76.3	8	21	AA1984 (F-18)
6	33	55.9	6	15	AA1985 (F-18)
7	30	50.8	8	12	AA1986 (F-18)
8	28	47.5	7	19	AA1987 (F-18)
9	28	47.5	8	21	AA1988 (F-18)
10	28	47.5	8	21	AA1989 (F-18)
11	28	47.5	8	22	AA1990 (F-18)

12	28	47.5	8	22	AB02392	Viral core polypep
13	28	47.5	8	22	AA13472	DPI78-like/DPI07-1
14	28	47.5	8	22	AA13472	Core polypeptide T
15	27	45.8	6	15	AA157812	RGD contg. peptide
16	27	45.8	6	15	AA157813	RGD contg. peptide
17	27	45.8	6	15	AA157801	RGD contg. peptide
18	27	45.8	6	15	AA157803	RGD contg. peptide
19	27	45.8	6	15	AA157803	G. oxydans cytochr
20	26	44.1	7	22	AA197759	Anti-b-endorphin m
21	26	44.1	6	13	AA120449	Beta-endorphin N-t
22	26	44.1	7	20	AA170043	Membrane dipeptida
23	26	44.1	8	15	AA160436	Antiproliferative
24	26	44.1	8	18	AA122683	BSW17 peptide mime
25	26	44.1	8	18	AA122684	BSW17 peptide mime
26	26	44.1	8	18	AA122684	Immunogenic peptid
27	25.5	43.2	8	21	AA127889	Anti-b-endorphin m
28	25	42.4	6	13	AA120437	Bovine Serum Album
29	25	42.4	6	15	AA155695	Snake venom derive
30	25	42.4	6	19	AA155695	Snake venom derive
31	25	42.4	6	19	AA155695	Snake venom derive
32	25	42.4	6	19	AA155695	Snake venom derive
33	25	42.4	6	19	AA155695	Snake venom derive
34	25	42.4	7	10	AA191775	Synthetic SCM-acti
35	25	42.4	7	10	AA191775	Synthetic SCM-acti
36	25	42.4	7	20	AA133390	A. diadematus fibr
37	25	42.4	7	22	AA190337	Silk spider fibrin
38	25	42.4	8	15	AA161040	Dynorphin-like pol
39	25	42.4	8	19	AA155695	Snake venom derive
40	25	42.4	8	19	AA155695	T-cell receptor CD
41	25	42.4	8	19	AA155695	Snake venom derive
42	25	42.4	8	19	AA155695	Snake venom derive
43	25	42.4	8	19	AA155695	Snake venom derive
44	25	42.4	8	21	AA133894	Human ART-1 peptid
45	25	42.4	8	21	AA188615	T-cell receptor co

ALIGNMENTS

RESULT 1

AA1903714
ID AA1903714 standard; peptide; 8 AA.

AC AA1903714;

DT 08-JUN-1999 (first entry)

DE Fluorine-18 (F-18) labeled peptide 1.

XX 18F radionuclide; targeting vector; positron emission tomography; F-18;
KW radiolabeling; thiol; fluorine-18.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "optionally has a free or protected
FT amino acid group"

FT Misc-difference 2 /note= "D-form residue"

FT Misc-difference 3 /note= "D-form residue"

FT Misc-difference 5 /note= "D-form residue"

FT Misc-difference 7 /note= "D-form residue; optionally has a free or
FT protected amino acid group"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

XX WO911590-A1.

XX 11-MAR-1999.

XX 11-MAR-1999.

```

XX 03-SEP-1998; 98WO-US18268.
XX
XX 03-SEP-1997; 97US-0057485.
XX
XX (IMMU-) IMMUNOMEDICS INC.
XX
XX Griffiths GL;
XX
XX WPI; 1999-228967/19.
XX
XX Radiolabeling thiol-containing peptides with fluorine-18
XX
XX Claim 13; Page 15; 22pp; English.
XX
XX The invention relates to a method for incorporating 18F radionuclide
XX into peptide-containing targeting vectors for use in clinical positron
XX emission tomography. Radiolabeling thiol-containing peptides with
XX fluorine-18 (F-18) comprises reacting a peptide comprising a free thiol
XX group with a labeling reagent of formula: 18F-(CH2)m-CR1R2-(CH2)n-X, or
XX a fluorinated alkene in which at least one of the two double bonded
XX carbon atoms bears at least one leaving group comprising I, Br, Cl,
XX azide, tosylate, mesylate, nosylate or triflate. n, m = 0-2; n+m = 0-2;
XX X = I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, maleimide
XX (optionally substituted by 1-2 alkyl) or 3-sulfomaleimide; R1, R2 = I,
XX Br, Cl, azide, tosylate, mesylate, nosylate, triflate, H, CONH2, COOH,
XX OH, sulfonic acid, tertiary amine, quaternary ammonium, alkyl
XX (optionally substituted by CONH2, COOH, OH, sulfonic acid, tertiary amine
XX or quaternary ammonium), COOR', CONR'2 or COR'; and R' = 1-6C alkyl or
XX phenyl. The method is used for radiolabeling peptide-containing
XX targeting vectors such as proteins, antibodies, antibody fragments and
XX receptor- targeted peptides for use in routine clinical positron
XX emission tomography. The method is simple and efficient. The method uses
XX the unique property of the free thiol groups which are rapidly alkylated
XX at neutral pH and moderate temperature. Sequences AAY03714-716 represent
XX examples of F-18 labeled peptides used in the method of detecting a
XX tissue.
XX
XX Sequence 8 AA;
XX
XX Query Match 100.0%; Score 59; DB 20; Length 8;
XX Best Local Similarity 100.0%; Pred. No. 6.4e+05;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 gywgkgyw 8
XX |
XX Db 1 gywgkgyw 8
XX
XX RESULT 2
XX AAY03716
XX ID AAY03716 standard; peptide; 8 AA.
XX
XX AC AAY03716;
XX
XX DT 08-JUN-1999 (first entry)
XX
XX Fluorine-18 (F-18) labeled peptide 3.
XX
XX 18F radionuclide; targeting vector; positron emission tomography; F-18;
XX radiolabeling; thiol; fluorine-18.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Modified-site 1
XX /note= "N-terminal acetylation"
XX
XX Misc-difference 2
XX /note= "D-form residue; D-iodo-Tyr"
XX
XX Misc-difference 3
XX /note= "D-form residue"
XX
XX Misc-difference 5
XX

```

```

FT Misc-difference 7 /note= "D-form residue; optionally acetylated"
FT
FT Misc-difference 8 /note= "D-form residue; D-iodo-Tyr"
FT
FT Misc-difference 8 /note= "D-form residue"
XX
XX WO9911590-A1.
XX
XX 11-MAR-1999.
XX
XX 03-SEP-1998; 98WO-US18268.
XX
XX 03-SEP-1997; 97US-0057485.
XX
XX (IMMU-) IMMUNOMEDICS INC.
XX
XX Griffiths GL;
XX
XX WPI; 1999-228967/19.
XX
XX Radiolabeling thiol-containing peptides with fluorine-18
XX
XX Claim 15; Page 15; 22pp; English.
XX
XX The invention relates to a method for incorporating 18F radionuclide
XX into peptide-containing targeting vectors for use in clinical positron
XX emission tomography. Radiolabeling thiol-containing peptides with
XX fluorine-18 (F-18) comprises reacting a peptide comprising a free thiol
XX group with a labeling reagent of formula: 18F-(CH2)m-CR1R2-(CH2)n-X, or
XX a fluorinated alkene in which at least one of the two double bonded
XX carbon atoms bears at least one leaving group comprising I, Br, Cl,
XX azide, tosylate, mesylate, nosylate or triflate. n, m = 0-2; n+m = 0-2;
XX X = I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, maleimide
XX (optionally substituted by 1-2 alkyl) or 3-sulfomaleimide; R1, R2 = I,
XX Br, Cl, azide, tosylate, mesylate, nosylate, triflate, H, CONH2, COOH,
XX OH, sulfonic acid, tertiary amine, quaternary ammonium, alkyl
XX (optionally substituted by CONH2, COOH, OH, sulfonic acid, tertiary amine
XX or quaternary ammonium), COOR', CONR'2 or COR'; and R' = 1-6C alkyl or
XX phenyl. The method is used for radiolabeling peptide-containing
XX receptor- targeted peptides for use in routine clinical positron
XX emission tomography. The method is simple and efficient. The method uses
XX the unique property of the free thiol groups which are rapidly alkylated
XX at neutral pH and moderate temperature. Sequences AAY03714-716 represent
XX examples of F-18 labeled peptides used in the method of detecting a
XX tissue.
XX
XX Sequence 8 AA;
XX
XX Query Match 100.0%; Score 59; DB 20; Length 8;
XX Best Local Similarity 100.0%; Pred. No. 6.4e+05;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 gywgkgyw 8
XX |
XX Db 1 gywgkgyw 8
XX
XX RESULT 3
XX AAY76818
XX ID AAY76818 standard; peptide; 8 AA.
XX
XX AC AAY76818;
XX
XX DT 28-APR-2000 (first entry)
XX
XX Immunogenic peptide for bi-specific antibody recognition.
XX
XX Immunogenic peptide; bi-specific antibody; diagnosis; immune response;
XX diseased tissue identification; therapy.
XX
XX Synthetic.
XX

```

XX Key Location/Qualifiers
 FH Modified-site 1 /note= "acetylated"
 FT Misc-difference 2 /note= "D-form Iodo-Tyr"
 FT Misc-difference 3 /note= "D-form residue"
 FT Misc-difference 5 /note= "acetylated D-form residue"
 FT Misc-difference 7 /note= "D-form Iodo-Tyr"
 FT Misc-difference 8 /note= "D-form residue"
 XX WO9966951-A2.
 XX 29-DEC-1999.
 XX 22-JUN-1999; 99WO-US13879.
 XX 22-JUN-1998; 98US-0090142.
 PR 14-OCT-1998; 98US-0104156.
 XX (IMMU-) IMMUNOMEDICS INC.
 XX Hansen HJ, Griffiths GL, Leung S, McBride WJ, Qu Z;
 PI WPI; 2000-160561/14.
 XX Bi-specific antibodies that bind specific target tissue and targeted
 conjugates -
 Claim 23; Page 61; 76pp; English.
 XX This sequence represents an immunogenic peptide for a bi-specific
 antibody. The invention relates to a method of treating or identifying
 diseased tissues in a patient comprising administering a bi-specific
 antibody (or fragment) having at least 1 arm (A) that specifically
 binds a targeted tissue and at least 1 arm (B) that specifically binds a
 targetable conjugate. The methods and bi-specific antibodies and fusion
 proteins are useful for pre-targeting methods of diagnosis and therapy.
 CC It is advantageous to raise bi-specific antibodies against a targetable
 conjugate that is capable of carrying at least 1 diagnostic or
 CC therapeutic agent. The characteristics of the chelator, metal chelate
 CC complex, therapeutic agent or diagnostic agent can be varied to
 CC accommodate differing applications without raising new bi-specific
 CC antibodies for each new application. The targetable conjugate is selected
 CC to elicit sufficient immune responses and also for rapid in vivo
 CC clearance when used within the bi-specific antibody targeting method.
 XX Sequence 8 AA;
 SQ
 Query Match 86.4%; Score 51; DB 21; Length 8;
 Best Local Similarity 87.5%; Pred. No. 6.4e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 gywkggyw 8
 ||| ||||
 Db 1 gywhkgyw 8
 RESULT 4
 AA03715
 ID AA03715 standard; peptide; 8 AA.
 XX
 AC AA03715;
 XX
 DT 08-JUN-1999 (first entry)
 XX
 DE Fluorine-18 (F-18) labeled peptide 2.
 XX

KW 18F radionuclide; targeting vector; positron emission tomography; F-18;
 KW radiolabeling; thiol; fluorine-18.
 XX Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 1 /note= "N-terminal acetylation; optionally has a free
 FT or protected thiol group"
 FT Misc-difference 2 /note= "D-form residue"
 FT Misc-difference 3 /note= "D-form residue"
 FT Misc-difference 5 /note= "D-form residue; optionally has a free or
 FT protected thiol group"
 FT Misc-difference 7 /note= "D-form residue"
 FT Misc-difference 8 /note= "D-form residue"
 FT WO9911590-A1.
 XX 11-MAR-1999.
 PD 03-SEP-1998; 98WO-US18268.
 XX 03-SEP-1997; 97US-0057485.
 XX (IMMU-) IMMUNOMEDICS INC.
 PA Griffiths GL;
 PI WPI; 1999-228967/19.
 DR Radiolabeling thiol-containing peptides with fluorine-18
 XX Claim 14; Page 15; 22pp; English.
 XX The invention relates to a method for incorporating 18F radionuclide
 CC into peptide-containing targeting vectors for use in clinical positron
 CC emission tomography. Radiolabeling thiol-containing peptides with
 CC fluorine-18 (F-18) comprises reacting a peptide comprising a free thiol
 CC group with a labeling reagent of formula: 18F-(CH2)m-CRIR2-(CH2)n-X, or
 CC a fluorinated alkene in which at least one of the two double bonded
 CC carbon atoms bears at least one leaving group comprising I, Br, Cl,
 CC azide, tosylate, mesylate, nosylate or triflate. n, m = 0-2; n+m = 0-2;
 CC x = I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, maleimide
 CC (optionally substituted by 1-2 alkyl) or 3-sulfomaleimide; R1, R2 = I,
 CC Br, Cl, azide, tosylate, mesylate, nosylate, triflate, H, CONH2, COOH,
 CC OH, sulfonic acid, tertiary amine, quaternary ammonium, alkyl
 CC (optionally substituted by CONH2, COOH, OH, sulfonic acid, tertiary amine
 CC or quaternary ammonium), COOR', CONR'2 or COR'; and R' = 1-6C alkyl or
 CC phenyl. The method is used for Radiolabeling peptide-containing
 CC targeting vectors such as proteins, antibodies, antibody fragments and
 CC receptor- targeted peptides for use in routine clinical positron
 CC emission tomography. The method is simple and efficient. The method uses
 CC the unique property of the free thiol groups which are rapidly alkylated
 CC at neutral pH and moderate temperature. Sequences AA03714-716 represent
 CC examples of F-18 labeled peptides used in the method of detecting a
 CC tissue.
 XX Sequence 8 AA;
 SQ

Query Match 76.3%; Score 45; DB 20; Length 8;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywgkgyw 8
 ||| ||||
 Db 2 ywgqgyw 8

RESULT 5
AA76817
ID AAY76817 standard; peptide; 8 AA.
XX
AC AAY76817;
XX
DT 28-APR-2000 (first entry)
XX
DE Immunogenic peptide for bi-specific antibody recognition.
XX
KW Immunogenic peptide; bi-specific antibody; diagnosis; immune response;
KW diseased tissue identification; therapy.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1 /note= "acetylated; modified with free amino acid group,
FT protected amino acid group, chelating agent or a
FT metal-chelate complex"
FT
FT Misc-difference 2 /note= "D-form residue"
FT
FT Misc-difference 3 /note= "D-form residue"
FT
FT Misc-difference 5 /note= "D-form residue; modified with free amino acid
FT group, protected amino acid group, chelating
FT agent or a metal-chelate complex"
FT
FT Misc-difference 7 /note= "D-form residue"
FT
FT Misc-difference 8 /note= "D-form residue"
FT
FT
FT WO9966951-A2.
XX
PN 29-DEC-1999.
XX
PD 22-JUN-1999; 99WO-US13879.
XX
PF 22-JUN-1998; 98US-0090142.
XX
PR 14-OCT-1998; 98US-0104156.
XX
XX (IMMU-) IMMUNOMEDICS INC.
XX
XX Hansen HJ, Griffiths GL, Leung S, McBride WJ, Qu Z;
PI WPI; 2000-160561/14.
XX
DR Bi-specific antibodies that bind specific target tissue and targeted
XX conjugates -
XX
XX Claim 22; Page 61; 76pp; English.
XX
XX This sequence represents an immunogenic peptide for a bi-specific
CC antibody. The invention relates to a method of treating or identifying
CC diseased tissues in a patient comprising administering a bi-specific
CC antibody (or fragment) having at least 1 arm (A) that specifically
CC binds a targeted tissue and at least 1 arm (B) that specifically binds a
CC targetable conjugate. The methods and bi-specific antibodies and fusion
CC proteins are useful for pre-targeting methods of diagnosis and therapy.
CC It is advantageous to raise bi-specific antibodies against a targetable
CC conjugate that is capable of carrying at least 1 diagnostic or
CC therapeutic agent. The characteristics of the chelator, metal chelate
CC complex, therapeutic agent or diagnostic agent can be varied to
CC accommodate differing applications without raising new bi-specific
CC antibodies for each new application. The targetable conjugate is selected
CC to elicit sufficient immune responses and also for rapid in vivo
CC clearance when used within the bi-specific antibody targeting method.
XX
XX Sequence 8 AA;

Query Match 76.3%; Score 45; DB 21; Length 8;
Best Local Similarity 85.7%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 ywkgwyw 8
Db 2 ywgcgyw 8
RESULT 6
AAR57814
ID AAR57814 standard; peptide; 6 AA.
XX
AC AAR57814;
XX
DT 03-OCT-1994 (first entry)
XX
DE RGD contg. peptide having antithrombotic activity.
XX
KW Platelet; antithrombotic; inhibitor; cell adhesion;
KW cancer; metastasis; extracorporeal circulation; coagulation;
KW anticoagulant; cell fusion.
XX
OS Synthetic.
XX
PN WO9405696-A.
XX
PD 17-MAR-1994.
XX
XX 07-SEP-1993; 93WO-JP01262.
PF
PR 07-SEP-1992; 92JP-0238624.
PR
PR 18-AUG-1993; 93JP-0203962.
XX
XX (YAWA) NIPPON STEEL CORP.
XX
XX Katada J, Sato Y;
XX WPI; 1994-101121/12.
DR
XX New peptides and their salts - used in platelet preps., as
PT antithrombotic agents and as inhibitors of e.g. cell adhesion,
PT cancer metastasis and extracorporeal circulation coagulation
XX
XX Disclosure; Page 52; 88pp; Japanese.
PS
XX The peptides given in AAR57801-28 have antithrombotic, anticoagulant
CC and cell fusion inhibiting effects. They may be used in
CC antithrombotic agents and as extracorporeal circulation
CC coagulation inhibitors, cell adhesion inhibitors, cancer
CC metastasis inhibitors and protecting agent in platelet
CC preparations.
CC (IC50 = 2.2 x 10⁻⁵ M).
XX
SQ Sequence 6 AA;

Query Match 55.9%; Score 33; DB 15; Length 6;
Best Local Similarity 66.7%; Pred. No. 6.4e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 wdkgyw 8
Db 1 wgrgdw 6
RESULT 7
AAR11895
ID AAR11895 standard; Protein; 8 AA.
XX
AC AAR11895;
XX
DT 19-JUL-1991 (first entry)

XX Lycium B peptide.
DE
XX
KW Lycium; angiotensin converting enzyme; renin; hypertension; cyclic.
XX
FH Key Location/Qualifiers
FT Modified-site 1..1
FT /label= OTHER
FT /note= "pyroglutamyl"
FT Modified-site 4..4
FT /label= OTHER
FT /note= "glycyl Ca"
FT Modified-site ..8
FT /label= OTHER
FT /note= "tryptophan indol"
XX
XX JP03081293-A.
XX
XX
XX PD 05-APR-1991.
XX
XX PF 23-AUG-1989; 89JP-0217000.
XX
XX PR 23-AUG-1989; 89JP-0217000.
XX
XX PA (NIME-) NIPPON MEKTRON KK.
XX
XX DR WPI; 1991-144847/20.
XX
XX PT Peptide(s) lyciumin A and B - have ACE and renin inhibiting
XX activity for treatment of hypertension
XX
XX PS Claim 1; page 1; 6pp; Japanese.
XX
XX CC This lycium B peptide has angiotensin converting enzyme (ACE) and
XX renin inhibiting activity and can be used in the treatment of
XX hypertension. It is used in the form of a drug compsn. Residue 4
XX (Gly) is modified by calcium and residue 8 (trp) has an indol gp.
XX present.
XX CC See also AAR11894.
XX
XX SQ Sequence 8 AA;

Query Match 50.8%; Score 30; DB 12; Length 8;
Best Local Similarity 66.7%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 wgkgyw 8
| | | |
Db 3 wgvgs 8

RESULT 8
AAW46276
ID AAW46276 standard; Peptide; 7 AA.
XX
XX AC AAW46276;
XX
XX DT 17-AUG-1998 (first entry)
XX
XX DE Moraxella catarrhalis CopB region I epitope.
XX
XX KW CopB; outer membrane protein; epitope; infection; diagnosis;
XX therapy; vaccine.
XX
XX OS Moraxella catarrhalis strain 035E.
XX Synthetic.
XX
XX PN WO9806851-A2.
XX
XX PD 19-FEB-1998.
XX
XX PF 12-AUG-1997; 97WO-US14221.
XX
XX KW CopB; outer membrane protein; epitope; infection; diagnosis;
XX therapy; vaccine.
XX
XX OS Moraxella catarrhalis strain 035E.
XX Synthetic.
XX
XX PN WO9806851-A2.
XX
XX PD 19-FEB-1998.
XX
XX PF 12-AUG-1997; 97WO-US14221.

XX 12-AUG-1996; 96US-0023832.
XX (TEXA) UNIV TEXAS SYSTEM.
XX
XX PI Aebi C, Cope LD, Hansen EJ;
XX WPI; 1998-159542/14.
XX
XX PT New isolated Moraxella catarrhalis peptide(s) - which define
XX epitopes of the outer membrane protein used to develop products for
XX the diagnosis, prophylaxis and treatment of infection
XX
XX PS Claim 3; Page 113; 132pp; English.
XX
XX CC This peptide corresponds to amino acid residues 296-302, in the
XX non-conserved region 1, of the outer membrane protein CopB (see
XX AAW46271) of Moraxella catarrhalis strain 035E. Claimed isolated
XX peptides of about 5-60 amino acid residues comprise at least
XX this 7-amino acid peptide sequence and are reactive with
XX monoclonal antibody 10F3. The 7-amino peptide is most preferred
XX for binding to 10F3. It is located C-terminal to the Asn-295
XX residue of CopB that influences the binding kinetics of the epitope.
XX Peptides which contain residues C-terminal to Asn-295 retain
XX reactivity, but this reactivity is less than that obtained in the
XX presence of Asn-295. Claimed peptides (see AAW46274-84) from
XX non-conserved region 1 can be used in the diagnosis, prophylaxis
XX (as vaccines) and treatment of M. catarrhalis infections.
XX
XX SQ Sequence 7 AA;

Query Match 47.5%; Score 28; DB 19; Length 7;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ywgkgy 7
| | | |
Db 2 yagkgy 7

RESULT 9
AAW46279
ID AAW46279 standard; Peptide; 8 AA.
XX
XX AC AAW46279;
XX
XX DT 17-AUG-1998 (first entry)
XX
XX DE Moraxella catarrhalis CopB region I epitope.
XX
XX KW CopB; outer membrane protein; epitope; infection; diagnosis;
XX therapy; vaccine.
XX
XX OS Moraxella catarrhalis strain 035E.
XX Synthetic.
XX
XX PN WO9806851-A2.
XX
XX PD 19-FEB-1998.
XX
XX PF 12-AUG-1997; 97WO-US14221.
XX
XX PR 12-AUG-1996; 96US-0023832.
XX
XX PA (TEXA) UNIV TEXAS SYSTEM.
XX
XX PI Aebi C, Cope LD, Hansen EJ;
XX WPI; 1998-159542/14.
XX
XX PT New isolated Moraxella catarrhalis peptide(s) - which define
XX epitopes of the outer membrane protein used to develop products for

PT the diagnosis, prophylaxis and treatment of infection
XX Claim 6; Page 113; 132pp; English.

XX This peptide corresponds to amino acid residues 295-302, in the
CC non-conserved region 1, of the outer membrane protein CopB (see
CC AAW46271) of Moraxella catarrhalis strain 035E. Claimed isolated
CC peptides of about 5-60 amino acid residues comprise at least
CC this 8-amino acid peptide sequence and are reactive with
CC monoclonal antibody 10F3. The 8-amino peptide is most preferred
CC for binding to 10F3. It includes the Asn-295 residue of CopB that
CC is required for 10F3 binding. Claimed peptides (see AAW46274-84)
CC from non-conserved region 1 can be used in the diagnosis,
CC prophylaxis (as vaccines) and treatment of M. catarrhalis
CC infections.

XX Sequence 8 AA;

Query Match 47.5%; Score 28; DB 19; Length 8;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 ywgkgy 7
| | | | |
Db 3 yagkgy 8

RESULT 10

AAW89518
ID AAW89518 standard; peptide; 8 AA.

XX AAW89518;

XX 23-MAY-2000 (first entry)

XX Core polypeptide fragment T No. 1076.

DE Retrovirus; hybrid polypeptide; enhancer; gp41; envelope protein; HIV-1;
XX HIV-2; SIV; pharmacokinetic; half-life; growth factor; cytokine; viral;
KW anti-fusogenic; differentiation factor; interleukin; interferon;
KW colony stimulating factor; hormone; angiogenic factor.

XX Unidentified.

OS WO9959615-A1.

XX 25-NOV-1999.

XX 20-MAY-1999; 99WO-US11219.

XX 20-MAY-1998; 98US-0082279.

XX (TRIM-) TRIMERIS INC.

XX Barney S, Guthrie KI, Merutka G, Anwer MK, Lambert DM;

XX WPI; 2000-136792/12.

XX A new hybrid polypeptide with enhanced pharmacokinetic properties
PT comprises enhancer sequence -

XX Disclosure; Page 39; 124pp; English.

XX The invention relates to hybrid polypeptides comprising enhancer peptide
CC sequence linked to core polypeptides. The enhancer polypeptides are
CC derived from various retroviral envelope (gp41) protein sequences,
CC especially from HIV-1, HIV-2 and SIV. The enhancer peptides enhance the
CC pharmacokinetic properties such as increasing the half-life of any core
CC polypeptide that they are linked to. The core polypeptides are any
CC polypeptide that may be introduced into a living system and that can
CC function as a pharmacologically useful peptide for the treatment or
CC prevention of a disease. The core polypeptides are bioactive peptides

CC selected from a growth factor, cytokine, differentiation factor,
CC interleukin, interferon, colony stimulating factor, hormone or
CC angiogenic factor. The peptides of the invention can be used for
CC inhibiting viral infection and can be used in anti-viral and
CC anti-fusogenic treatments. Sequences AAW8651-Y90055 represent core
CC polypeptide fragments that can be used in the invention. Some sequences
CC among those indicated also comprise enhancer fragments at terminal ends
XX and form hybrid polypeptides.

SQ Sequence 8 AA;

Query Match 47.5%; Score 28; DB 21; Length 8;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 wqkgy 7
| | | | |
Db 3 wgygy 7

RESULT 11

ABB00926

ID ABB00926 standard; Peptide; 8 AA.

XX ABB00926;

XX 03-JAN-2002 (first entry)

XX Viral DP178/107-like region peptide T1076.

DE Human immunodeficiency virus; HIV; respiratory syncytial virus; RSV;
XX virucide; heptad repeat region; transmembrane protein; gp41; HR1; HR2;
KW infection.

XX Virididae.

XX Key

FT Modified-site 1 Location/Qualifiers

FT /note= "N-terminal is substituted by Ac"

FT Modified-site 8

FT /note= "C-terminal amide"

PN WO200164013-A2.

XX 07-SEP-2001.

XX 07-FEB-2001; 2001WO-US03988.

XX 29-FEB-2000; 2000US-0515965.

XX (TRIM-) TRIMERIS INC.

XX Antczak JB, Delmedico MK, Erickson JB, Lambert DM, Sista P;

XX WPI; 2001-514829/56.

XX Heptad repeat region peptide analogs useful for inhibiting virus/cells
PT fusion, useful for treating HIV and Respiratory Syncytial Virus
PT infection -

XX Disclosure; Page 51; 587pp; English.

XX The invention relates to isolated analogues of the heptad repeat region
CC peptides DP178 and DP107. DP178 and DP107 correspond to amino acids
CC 638-673 (heptad repeat region HR2) and 558-595 (heptad repeat region
CC HR1) respectively, of HIV-1/IIA1 transmembrane protein gp41. The HR1
CC and HR2 regions of proteins interact non-covalently with each other
CC and/or with peptides derived from them. This interaction is required for
CC normal infectivity of viruses such as RSV and HIV. The heptad
CC repeat region peptide analogues may be used to inhibit respiratory
CC syncytial virus (RSV) infection in a cell. They may also be used to
CC inhibit HIV infection. The present sequence is a peptide provided in

CC the specification.
XX
SQ Sequence 8 AA;

Query Match 47.5%; Score 28; DB 22; Length 8;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 wgkgy 7
|||
Db 3 wgygy 7

RESULT 12
ABB02392
ID ABB02392 standard; Peptide; 8 AA.
XX
AC ABB02392;
XX
DT 03-JAN-2002 (first entry)
XX
DE Viral core polypeptide, SEQ ID NO: 919.
XX
KW Human immunodeficiency virus; HIV; respiratory syncytial virus; RSV;
KW virucide; heptad repeat region; transmembrane protein; gp41; HR1; HR2;
KW infection.
XX
OS Virididae.
XX
PN WO200164013-A2.
XX
PD 07-SEP-2001.
XX
PF 07-FEB-2001; 2001WO-US03988.
XX
PR 29-FEB-2000; 2000US-0515965.
XX
PA (TRIM-) TRIMERIS INC.
XX
PI Antczak JB, Delmedico MK, Erickson JB, Lambert DM, Sista P;
XX WPI; 2001-514829/56.
XX
DR Heptad repeat region peptide analogs useful for inhibiting virus/cells
PT fusion, useful for treating HIV and Respiratory Syncytial Virus
PT infection -
XX
PS Disclosure; Page 410; 587pp; English.

The invention relates to isolated analogues of the heptad repeat region peptides DP178 and DP107. DP178 and DP107 correspond to amino acids 638-673 (heptad repeat region HR2) and 558-595 (heptad repeat region HR1) respectively, of HIV-1LAI transmembrane protein gp41. The HR1 and HR2 regions of proteins interact non-covalently with each other and/or with peptides derived from them. This interaction is required for normal infectivity of viruses such as RSV and HIV. The heptad repeat region peptide analogues may be used to inhibit respiratory syncytial virus (RSV) infection in a cell. They may also be used to inhibit HIV infection. The present sequence is a peptide provided in the specification.

XX
SQ Sequence 8 AA;

Query Match 47.5%; Score 28; DB 22; Length 8;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 wgkgy 7
|||
Db 3 wgygy 7

RESULT 13
AAU13472
XX
AC AAU13472;
XX
DT 21-NOV-2001 (first entry)
XX
DE DP178-like/DP107-like peptide T-1076.
XX
DE Anti-retroviral; DP178-like; DP107-like; transmembrane protein gp41;
KW antifusogenic; antiviral; HIV transmission; mutant; mutcin.
XX
KW Human immunodeficiency virus 1 isolate LAI.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal is substituted by Ac"
FT Modified-site 8
FT /note= "C-terminal amide"
FT
XX
PN WO200151673-A2.
XX
PD 19-JUL-2001.
XX
PF 05-JUL-2000; 2000WO-US35727.
XX
PR 09-JUL-1999; 99US-0350841.
XX (TRIM-) TRIMERIS INC.
XX
PI Jeffs P, Lackey JW, Erickson JB, Lawless MK, Merutka G;
XX WPI; 2001-442157/47.
XX
DR Identifying a compound that inhibits the formation of or disrupts a
PT DP107/DP178 complex, especially compounds with antifusogenic, antiviral
PT or intracellular modulatory activity, by detecting the formation of a
PT DP107/DP178 complex -
XX
PS Disclosure; Page 71; 259pp; English.

The present invention relates to peptides which exhibit anti-retroviral activity. The peptides of the invention (AAU12559-AAU14009) comprise DP178-like and DP107-like peptides. The DP178 peptide corresponds to amino acids 639-673 of the transmembrane protein gp41 from human immunodeficiency virus 1 (HIV-1) isolate LAI. The DP107 peptide corresponds to amino acids 558-595 of gp41 from HIV-1LAI. The invention also relates to a method of identifying compounds that inhibit the formation of or disrupts a DP107/DP178 complex. The method comprises detecting the formation of a DP107/DP178 complex, both in the presence or absence of a test compound, in a reaction mixture containing DP107 and DP178 peptides. The method is useful for identifying compounds, including small molecule compounds, which may themselves exhibit antifusogenic, antiviral or intracellular modulatory activity. The DP178-like/DP107-like peptides are useful to inhibit human and non-human retroviral, particularly HIV, transmission to uninfected cells. The present sequence represents one of the DP178-like/DP107-like peptides of the invention.

XX
SQ Sequence 8 AA;

Query Match 47.5%; Score 28; DB 22; Length 8;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 wgkgy 7
|||
Db 3 wgygy 7

RESULT 14

AAB77919
ID AAB77919 standard; Peptide; 8 AA.XX AC
XX AC AAB77919;

DT 19-APR-2001 (first entry)

XX DE Core polypeptide T1076.

XX KW Core polypeptide; enhancer; antiviral; anti-HIV;
KW virucide; hepatotropic; antiinflammatory; hybrid polypeptide;
KW coiled-coil peptide interaction; fusion-related disorder;
KW bacterial infection; viral infection.

XX OS Unidentified.

XX PN WO200103723-A1.

XX PD 18-JAN-2001.

XX PF 10-JUL-2000; 2000WO-US18772.

XX PR 09-JUL-1999; 99US-0350641.

XX PA (TRIM-) TRIMERIS INC.

XX PI Barney S, Guthrie KI, Merutka G, Anwer MK, Lambert DM;

XX DR WPI; 2001-147136/15.

XX PT New hybrid polypeptide, useful for preventing, treating and diagnosing
e.g. viral infections, comprises an enhancer peptide linked to a core
polypeptide -

XX PS Disclosure; Page 51; 151pp; English.

The present sequence is a core polypeptide which may be linked to an enhancer peptide to form a novel hybrid polypeptide. The hybrid polypeptide exhibits enhanced pharmacokinetic properties relative to those exhibited by the core polypeptide when introduced into a living system. It is used to increase the in vitro or ex vivo half-life of the core polypeptide. The hybrid and core polypeptides can be used for modulating fusogenic events and intracellular processes involving coiled-coil peptide interactions. Other uses include preventing, treating and/or diagnosing disorders involving fusion events (e.g. modulation of neurotransmitter exchange and sperm-egg fusion), intracellular processes involving coiled-coil peptides (e.g. bacterial infections) and viral infections that involve cell-cell and/or virus-cell fusion (e.g. viral infections caused by human immunodeficiency virus, respiratory syncytial virus, Epstein-Barr virus, hepatitis B virus, Mason-Pfizer virus and polio virus). The enhancer peptide sequence increases the half-life and reduces the clearance rate of therapeutic peptides, which increases their efficacy and minimises the incidence and severity of adverse side effects. In addition, this increases the sensitivity of the diagnostic procedure in which they are used.

XX SQ Sequence 8 AA;

Query Match 47.5%; Score 28; DB 22; Length 8;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 wqkgy 7

DB ||||

3 wgygy 7

RESULT 15

AAR57812

ID AAR57812 standard; peptide; 6 AA.

XX AC AAR57812;

XX DT 03-OCT-1994 (first entry)

XX DE RGD contg. peptide having antithrombotic activity.

XX KW Platelet; antithrombotic; inhibitor; cell adhesion;
KW cancer; metastasis; extracorporeal circulation; coagulation;
KW anticoagulant; cell fusion.

XX OS Synthetic.

XX PN WO9405696-A.

XX PD 17-MAR-1994.

XX PF 07-SEP-1993; 93WO-JP01262.

XX PR 07-SEP-1992; 92JP-0238624.

XX PR 18-AUG-1993; 93JP-0203962.

XX PA (YAWA) NIPPON STEEL CORP.

XX PI Katada J, Sato Y;

XX DR WPI; 1994-101121/12.

XX PT New peptides and their salts - used in platelet preps., as
antithrombotic agents and as inhibitors of e.g. cell adhesion,
cancer metastasis and extracorporeal circulation coagulation

XX PS Disclosure; Page 52; 88pp; Japanese.

The peptides given in AAR57801-28 have antithrombotic, anticoagulant and cell fusion inhibiting effects. They may be used in antithrombotic agents and as extracorporeal circulation coagulation inhibitors, cell adhesion inhibitors, cancer metastasis inhibitors and protecting agent in platelet preparations.
CC (IC50 = 3.1 x 10⁻⁵ M).

XX SQ Sequence 6 AA;

Query Match 45.8%; Score 27; DB 15; Length 6;
Best Local Similarity 50.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 wqkgyw 8

DB | : | |

1 wargdw 6

Search completed: August 12, 2002, 17:14:32
Job time: 245 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:12:17 ; Search time 20.43 seconds
(without alignments)
9.565 Million cell updates/sec

Title: 10-071247-1
Perfect score: 59
Sequence: 1 gywgkgyw 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 48605

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	49.2	8	2	US-08-467-046-22
2	28	47.5	4	1	US-08-358-160-227
3	28	47.5	8	4	US-09-082-279B-919
4	28	47.5	8	4	US-09-315-304B-919
5	26	44.1	6	1	US-07-718-577-2
6	26	44.1	7	4	US-09-258-754-9
7	26	44.1	7	4	US-09-042-107-9
8	26	44.1	8	1	US-08-014-426-28
9	26	44.1	8	5	PCT-US94-01319-28
10	25	42.4	6	1	US-08-232-261-9
11	25	42.4	6	6	US-08-462-661A-38
12	25	42.4	6	6	5318899-55
13	25	42.4	8	1	US-08-462-661A-19
14	25	42.4	8	4	US-08-525-539A-39
15	25	42.4	8	6	5318899-41
16	24	40.7	4	1	US-08-358-160-225
17	24	40.7	4	1	US-08-358-160-226
18	24	40.7	5	1	US-08-353-400-27
19	24	40.7	5	6	5185431-15
20	24	40.7	6	1	US-07-718-577-6
21	24	40.7	6	2	US-08-350-260A-14
22	24	40.7	7	1	US-08-487-006-111
23	24	40.7	7	2	US-08-463-230A-4
24	24	40.7	7	2	US-08-463-230A-5
25	24	40.7	7	2	US-08-488-659A-111
26	24	40.7	7	3	US-09-129-075-16
27	24	40.7	8	1	US-08-233-558-1

28 24 40.7 8 1 US-08-233-558-10
29 24 40.7 8 1 US-08-233-558-25
30 24 40.7 8 1 US-08-249-371-7
31 24 40.7 8 4 US-08-160-604-28
32 24 40.7 8 4 US-08-160-604-29
33 24 40.7 8 4 US-08-586-670A-17
34 24 40.7 8 4 US-09-537-357-25
35 24 40.7 8 5 PCT-US95-06451-7
36 23 39.0 6 1 US-08-321-625-27
37 23 39.0 6 1 US-08-321-625-28
38 23 39.0 6 4 US-09-181-083-27
39 23 39.0 6 4 US-09-181-083-28
40 23 39.0 7 4 US-09-258-754-167
41 23 39.0 7 4 US-09-042-107-167
42 23 39.0 8 1 US-07-938-334C-22
43 23 39.0 8 4 US-08-444-818-512
44 23 39.0 8 4 US-08-444-818-513
45 23 39.0 8 4 US-08-444-818-514

ALIGNMENTS

RESULT 1
US-08-467-046-22
: Sequence 22, Application US/08467046
: Patent No. 5948644
: GENERAL INFORMATION:
: APPLICANT: DOPHEIDE, THEODORUS AA
: APPLICANT: FRENKEL, MAURICE J
: APPLICANT: GRANT, WARWICK N
: APPLICANT: SAVIN, KEITH W
: APPLICANT: WAGLAND, BARRY M
: TITLE OF INVENTION: VACCINE
: NUMBER OF SEQUENCES: 32
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FOLEY & LARDNER
: STREET: SUITE 500, 1800 DIAGONAL ROAD
: CITY: ALEXANDRIA
: STATE: VIRGINIA
: COUNTRY: UNITED STATES OF AMERICA
: ZIP: 22313-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/467,046
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/904,055
: FILING DATE: 26-JUN-1992
: APPLICATION NUMBER: AU PJ0621
: FILING DATE: 26-SEP-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: AU PJ0622
: FILING DATE: 26-SEP-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: AU PJ0623
: FILING DATE: 26-SEP-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: AU PJ0624
: FILING DATE: 26-SEP-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/AU89/00416
: FILING DATE: 26-SEP-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/548,901
: FILING DATE: 26-JUL-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, STEPHEN A

; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/166
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Trichostrongylus colubriformis
US-08-467-046-22

Query Match 49.2%; Score 29; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 wgkyw 8
| | | |
Db 1 WMKGQW 6

RESULT 2
US-08-358-160-227
; Sequence 227, Application US/08358160
; Patent No. 5663143
; GENERAL INFORMATION:
; APPLICANT: LEY, Arthur C.
; APPLICANT: LADNER, Robert C.
; APPLICANT: GUTERMAN, Sonia K.
; APPLICANT: ROBERTS, Bruce L.
; APPLICANT: MARKLAND, William
; APPLICANT: KENT, Rachel B.
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; TITLE OF INVENTION: DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
; NUMBER OF SEQUENCES: 234
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W. Suite 300
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/358,160
; FILING DATE: 16-DEC-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/133,031
; FILING DATE: 13-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,319
; FILING DATE: 26-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/664,989
; FILING DATE: 01-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/487,063
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/240,160
; FILING DATE: 02-SEP-1988
; ATTORNEY/AGENT INFORMATION:

; NAME: Cooper, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: LEY=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 227:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-358-160-227

Query Match 47.5%; Score 28; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 wgkg 6
| | | |
Db 1 WKGK 4

RESULT 3
US-09-082-279B-919
; Sequence 919, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; TITLE OF INVENTION: PHARMACOKINETIC PROPERTIES
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 919
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-919

Query Match 47.5%; Score 28; DB 4; Length 8;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 wgkyg 7
| | | |
Db 3 WGYGY 7

RESULT 4
US-09-315-304B-919
; Sequence 919, Application US/09315304B
; Patent No. 6348568
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; TITLE OF INVENTION: PROPERTIES

; FILE REFERENCE: 7872-052
; CURRENT APPLICATION NUMBER: US/09/315,304B
; CURRENT FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1667
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 919
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-315-304B-919

Query Match 47.5%; Score 28; DB 4; Length 8;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 wkgky 7
|||
Db 3 WGYGY 7

RESULT 5
US-07-718-577-2
; Sequence 2, Application US/07718577
; Patent No. 5432018
; GENERAL INFORMATION:
; APPLICANT: Dower, William J.
; APPLICANT: Cwirla, Steven E.
; APPLICANT: Barrett, Ronald W.
; TITLE OF INVENTION: PEPTIDE LIBRARY AND
; TITLE OF INVENTION: SCREENING SYSTEMS
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Street
; STREET: Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07718,577
; FILING DATE: 19910620
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/541,108
; FILING DATE: 20-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11509-25-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-718-577-2

Query Match 44.1%; Score 26; DB 1; Length 6;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 gywg 4
|:
Db 2 GFWG 5

RESULT 6
US-09-258-754-9
; Sequence 9, Application US/09258754
; Patent No. 6174687
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; APPLICANT: Rajotte, Daniel
; TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using
; TITLE OF INVENTION: Membrane Diptidase
; FILE REFERENCE: P-LJ 3443
; CURRENT APPLICATION NUMBER: US/09/258,754
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/042,107
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 452
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-258-754-9

Query Match 44.1%; Score 26; DB 4; Length 7;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gywgk 5
|:
Db 3 GKWK 7

RESULT 7
US-09-042-107-9
; Sequence 9, Application US/09042107
; Patent No. 6232287
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Molecules that Home to Various Selected Organs or
; TITLE OF INVENTION: Tissues
; FILE REFERENCE: P-LJ 2892
; CURRENT APPLICATION NUMBER: US/09/042,107
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 436
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-042-107-9

Query Match 44.1%; Score 26; DB 4; Length 7;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gywgk 5
|:
Db 3 GKWK 7

Db 3 GKWK 7

```
RESULT 8
US-08-014-426-28
; Sequence 28, Application US/08014426
; Patent No. 5512435
; GENERAL INFORMATION:
; APPLICANT: Renschler, Markus F.
; APPLICANT: Levy, Ronald
; APPLICANT: Bhatt, Ramesh
; APPLICANT: Dower, William
; TITLE OF INVENTION: RECEPTOR-BINDING ANTIPROLIFERATIVE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,426
; FILING DATE: 05-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 5490A-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
US-08-014-426-28

Query Match 44.1%; Score 26; DB 1; Length 8;
Best Local Similarity 42.9%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ywgkgyw 8
:| |
Db 1 HWNEYW 7

RESULT 9
PCT-US94-01319-28
; Sequence 28, Application PC/TUS9401319
; GENERAL INFORMATION:
; APPLICANT: Renschler, Markus F.
; APPLICANT: Levy, Ronald
; APPLICANT: Bhatt, Ramesh
; APPLICANT: Dower, William
; TITLE OF INVENTION: RECEPTOR-BINDING ANTIPROLIFERATIVE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
```

```
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01319
; FILING DATE: 04-FEB-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/014,426
; FILING DATE: 05-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 5490A-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
PCT-US94-01319-28

Query Match 44.1%; Score 26; DB 5; Length 8;
Best Local Similarity 42.9%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ywgkgyw 8
:| |
Db 1 HWNEYW 7

RESULT 10
US-08-232-261-9
; Sequence 9, Application US/08232261
; Patent No. 5498601
; GENERAL INFORMATION:
; APPLICANT: Sato, Yoshimi
; APPLICANT: Hayashi, Yoshio
; APPLICANT: Katada, Jun
; TITLE OF INVENTION: No. 5498601el Peptides, and Platelet
; TITLE OF INVENTION: Aggregation-Inhibiting Agents, Blood
; TITLE OF INVENTION: Coagulation-Inhibiting Agents for Extracorporeal
; TITLE OF INVENTION: Circulation, Cell Adhesion-Inhibiting Agents, Tumor
; TITLE OF INVENTION: Metastasis-Inhibiting Agents...
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mr. Edward W. Greason, Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,261
; FILING DATE: 06-MAY-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
```



```
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD L. NEELEY, Ph.D.
; REGISTRATION NUMBER: 30,092
; REFERENCE/DOCKET NUMBER: CORT-003/08US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)843-5000
; TELEFAX: (415)857-0663
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 8
; OTHER INFORMATION: /note="this position is C-NH2."
;
US-08-462-661A-19

Query Match 42.4%; Score 25; DB 1; Length 8;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 gkgyw 8
| | | |
Db 2 GKGDW 6

RESULT 14
US-08-525-539A-39
; Sequence 39, Application US/08525539A
; Patent No. 6309636
; GENERAL INFORMATION:
; APPLICANT: DO COUTO, FERNANDO J.R.
; APPLICANT: CERIANI, ROBERTO L.
; APPLICANT: PETERSON, JERRY A.
; TITLE OF INVENTION: RECOMBINANT PEPTIDES DERIVED FROM THE
; TITLE OF INVENTION: NC3 ANTI-BA46 ANTIBODY, METHODS OF USE THEREOF, AND
; TITLE OF INVENTION: METHODS OF HUMANIZING ANTIBODY PEPTIDES
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1019
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,539A
; FILING DATE: 14-SEP-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: DYLAN, TYLER
; REGISTRATION NUMBER: 37,612
; REFERENCE/DOCKET NUMBER: 27633-20001.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-525-539A-39

; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD L. NEELEY, Ph.D.
; REGISTRATION NUMBER: 30,092
; REFERENCE/DOCKET NUMBER: CORT-003/08US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)843-5000
; TELEFAX: (415)857-0663
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 8
; OTHER INFORMATION: /note="this position is C-NH2."
;
US-08-462-661A-19

Query Match 42.4%; Score 25; DB 4; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 wgkgyw 8
| | |
Db 1 WGDYW 6

RESULT 15
5318899-41
; Patent No. 5318899
; APPLICANT: SCARBOROUGH, ROBERT M.; WOLF, DAVID L.; CHARO,
; ISRAEL F.
; TITLE OF INVENTION: PLATELET AGGREGATION INHIBITORS
; NUMBER OF SEQUENCES: 83
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/483,229
; FILING DATE: 20-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 418,028
; FILING DATE: 06-OCT-1989
; APPLICATION NUMBER: 367,509
; FILING DATE: 16-JUN-1989
; SEQ ID NO: 41
; LENGTH: 8
;
5318899-41

Query Match 42.4%; Score 25; DB 6; Length 8;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 gkgyw 8
| | | |
Db 2 GKGDW 6

Search completed: August 12, 2002, 17:14:59
Job time: 162 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:18:51 ; Search time 24.99 Seconds
(without alignments)
30.761 Million cell updates/sec

Title: 10-071247-2
Perfect score: 66
Sequence: 1 cywgcgyw 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 603

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	23	34.8	5	4	I79564
2	20	30.3	6	2	JH0253
3	19	28.8	6	2	F41946
4	19	28.8	8	2	S19288
5	18	27.3	6	2	A41946
6	18	27.3	7	2	B33882
7	18	27.3	7	2	S33567
8	17	25.8	4	2	B53284
9	17	25.8	6	2	A61068
10	17	25.8	6	2	PT0629
11	17	25.8	6	2	PT0637
12	17	25.8	7	2	S57274
13	17	25.8	7	2	PT0628
14	17	25.8	7	2	PT0642
15	17	25.8	7	2	PT0722
16	17	25.8	7	2	PT0586
17	17	25.8	7	2	PT0728
18	17	25.8	7	2	B48394
19	17	25.8	7	2	PD0029
20	17	25.8	8	2	A31570
21	17	25.8	8	2	JS0315
22	17	25.8	8	2	JS0316
23	17	25.8	8	2	JS0317
24	17	25.8	8	2	JS0318
25	17	25.8	8	2	PT0724
26	16	24.2	3	3	A22565
27	16	24.2	5	2	F22565
28	15	22.7	5	2	A33882
29	15	22.7	5	2	S65726

30	15	22.7	6	2	H48394	glycoprotein compo
31	15	22.7	6	2	PT0532	T-cell receptor be
32	15	22.7	7	2	A58512	venom heptapeptide
33	14	21.2	5	2	PT0278	Ig heavy chain CRD
34	14	21.2	6	2	B35640	cerebellar degener
35	14	21.2	7	2	S21230	dermorphin (irp-4,
36	14	21.2	7	2	S38516	dermorphin (irp-4,
37	14	21.2	7	2	E33932	mablinin II chain
38	14	21.2	8	2	A41117	Ig mu chain D regi
39	13	19.7	4	2	S09478	acetylcholinestera
40	13	19.7	6	2	A61411	globulin IV alpha
41	13	19.7	7	1	A61324	amelotin - rat
42	13	19.7	7	1	A61324	dermorphin - Rohde
43	13	19.7	7	2	A60139	fatty-acid synthas
44	13	19.7	7	2	S71870	glutathione transf
45	13	19.7	8	2	S36662	dermorphin (Lys-7)
						metallothionein is

ALIGNMENTS

RESULT 1

I79564

hypothetical TCL3 protein (mistranslated) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 28-Jan-2000 #sequence_revision 28-Jan-2000 #text_change 28-Jan-2000

C:Accession: I79564

R;Zutter, M.; Hockett, R.D.; Roberts, C.W.; McGuire, E.A.; Bloomstone, J.; Morton, C.,

Proc. Natl. Acad. Sci. U.S.A. 87, 3161-3165, 1990

A:Title: The t(10;14)(q24;q11) of T-cell acute lymphoblastic leukemia juxtaposes the

A:Reference number: I59162; MUID:90222189

A:Accession: I79564

A:Status: translation not shown; translated from GB/EMBL/DBBJ

A:Molecule type: DNA

A:Residues: 1-6 <ZUT>

A:Cross-references: GB:M33602; NID:g339907; PIDN:AAA66449.1; PID:g807656

C:Comment: This is the hypothetical translation of a sequence translated in an incorr

Query Match 34.8%; Score 23; DB 4; Length 6;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 wgcg 6

Db 1 1 1

2 WCCG 5

RESULT 2

JH0253

gut pentapeptide - Japanese eel

C:Species: Anguilla japonica (Japanese eel)

C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Apr-1995

C:Accession: JH0253

R;Uesaka, T.; Ikeda, T.; Kubota, I.; Muneoka, Y.;

Biochem. Biophys. Res. Commun. 180, 828-832, 1991

A:Title: Structure and function of a pentapeptide isolated from the gut of the eel.

A:Reference number: JH0253; MUID:92062113

A:Accession: JH0253

A:Molecule type: protein

A:Residues: 1-5 <UES>

A:Experimental source: gut

C:Comment: This peptide increased basal tone of the circular muscle of the esophagoga

, and of the circular muscle of the gastro-intestinal junction.

Query Match

Best Local Similarity

Matches

2; Conservative

1; Mismatches

0; Indels

0; Gaps

0;

30.3%; Score 20; DB 2; Length 5;

66.7%; Pred. No. 2.8e+05;

6 gw 8

1 1 1

1 GFW 3

```

RESULT 3
F41946
T-cell receptor gamma chain (1a.27) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C:Accession: F41946
R:Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma gene
A:Reference number: A41946; MUID:92049316
A:Accession: F41946
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-6 <WHE>
C:Keywords: T-cell receptor

Query Match 28.8%; Score 19; DB 2; Length 6;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 cgyw 8
| |
Db 2 CAVW 5

RESULT 4
S19288
acylase - Kluyvera cryocrescens
C:Species: Kluyvera cryocrescens
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
C:Accession: S19288
R:Martin, J.; Slade, A.; Aitken, A.; Arche, R.; Virden, R.
Biochem. J. 280, 659-662, 1991
A:Title: Chemical modification of serine at the active site of penicillin acylase from K
A:Reference number: S19288; MUID:92109664
A:Accession: S19288
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <MAR>

Query Match 28.8%; Score 19; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 cgyw 8
| |
Db 1 CNMW 4

RESULT 5
A41946
T-cell receptor gamma chain (1t.22) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C:Accession: A41946
R:Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma gene
A:Reference number: A41946; MUID:92049316
A:Accession: A41946
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-6 <WHE>
C:Keywords: T-cell receptor

Query Match 27.3%; Score 18; DB 2; Length 6;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 cgyw 8
| |
Db 1 CNMW 4

RESULT 6
B33882
cadmium-binding heptapeptide - downy thornapple
C:Species: Datura innoxia (downy thornapple)
C:Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 18-Jun-1993
C:Accession: B33882
R:Jackson, P.J.; Unkefer, C.J.; Doolen, J.A.; Watt, K.; Robinson, N.J.
Proc. Natl. Acad. Sci. U.S.A. 84, 6619-6623, 1987
A:Title: Poly(gamma-glutamylcysteinyl)glycine: its role in cadmium resistance in plan
A:Reference number: A94182; MUID:88016144
A:Accession: B33882
A:Molecule type: protein
A:Residues: 1-7 <QAJ>

Query Match 27.3%; Score 18; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 cywgcg 6
| |
Db 2 CECECG 7

RESULT 7
S33567
tubulin beta-3 chain - fruit fly (Drosophila melanogaster) (fragment)
C:Species: Drosophila melanogaster
C:Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 21-Jul-2000
C:Accession: S33567
R:Hinz, U.; Wolk, A.; Renkawitz-Pohl, R.
Development 110, 543-554, 1992
A:Title: Ultrabithorax is a regulator of beta-3 tubulin expression in the Drosophila
A:Reference number: S33567; MUID:93170162
A:Accession: S33567
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-7 <HIN>
A:Cross-references: EMBL:X68393; NID:g7669; PIDN:CAA48459.1; PID:g4377448
C:Genetics:
A:Gene: FlyBase:fbgr
A:Cross-references: FlyBase:FBgn0003888
A:Introns: 5/3

Query Match 27.3%; Score 18; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 gcgyw 8
| |
Db 3 GAKFW 7

RESULT 8
B53284
T-cell receptor beta 2 chain D region, Dbeta2 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: B53284
R:Harindranath, N.; Alexander, C.B.; Mage, R.G.
Mol. Immunol. 28, 881-888, 1991
A:Title: Evolutionarily conserved organization and sequences of germline diversity an
A:Reference number: A53284; MUID:91342695
A:Accession: B53284
A:Status: preliminary
A:Molecule type: DNA
```

A:Residues: 1-4 <HAR>

A:Cross-references: GB:S60737; NID:g233916; PIDN:AA19518.1; PID:g233918

A:Note: sequence extracted from NCBI backbone (NCBIN:60737, NCBIP:60738)

C:Keywords: T-cell receptor

Query Match 25.8%; Score 17; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 wg 4

Db 2 WG 3

RESULT 9

A61068

locustakinin - migratory locust

C:Species: Locusta migratoria (migratory locust)

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 08-Dec-1995

C:Accession: A61068

R:Schoofs, L.; Holman, G.M.; Proost, P.; Van Damme, J.; Hayes, T.K.; De Loof, A.

Regul. Pept. 37, 49-57, 1992

A:Title: Locustakinin, a novel myotropic peptide from Locusta migratoria, isolation, pri

A:Reference number: A61068; MUID:92262851

A:Accession: A61068

A:Molecule type: protein

A:Residues: 1-6 <SCH>

C:Keywords: amidated carboxyl end; cephalomytropic peptide; neuropeptide

F:6/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 25.8%; Score 17; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 wg 4

Db 5 WG 6

RESULT 10

PT0629

T-cell receptor beta chain V-D-J region (100-2AH) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C:Accession: PT0629; PT0528

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; MUID:91277601

A:Accession: PT0629

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-6 <FEE>

A:Experimental source: newborn thymus, strain BALB/c, clone 100-2AH

A:Accession: PT0528

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-6 <FE2>

A:Experimental source: adult thymus, strain BALB/c, clone 100-4AB

C:Keywords: T-cell receptor

Query Match 25.8%; Score 17; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 wg 4

Db 5 WG 6

RESULT 11

PT0637

T-cell receptor beta chain V-D-J region (111-1K) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C:Accession: PT0637

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions

A:Reference number: PT0509; MUID:91277601

A:Accession: PT0637

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-6 <FEE>

A:Experimental source: newborn thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 25.8%; Score 17; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 wg 4

Db 5 WG 6

RESULT 12

S57274

triacylglycerol lipase (EC 3.1.1.3) - Psychrobacter immobilis (fragment)

C:Species: Psychrobacter immobilis

C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 03-Nov-1995

C:Accession: S57274

R:Arpigny, J.L.; Feller, G.; Gerday, C.

Biochim. Biophys. Acta 1263, 103, 1995

A:Title: Corrigendum to "Cloning, sequence and structural features of a lipase from t

A:Reference number: S57274; MUID:95359197

A:Accession: S57274

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-7 <ARP>

A:Cross-references: EMBL:X67712

C:Keywords: carboxylic ester hydrolase

Query Match 25.8%; Score 17; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 wg 4

Db 1 WG 2

RESULT 13

PT0628

T-cell receptor beta chain V-D-J region (111-1AG) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C:Accession: PT0628

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions

A:Reference number: PT0509; MUID:91277601

A:Accession: PT0628

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-7 <FEE>

A:Experimental source: newborn thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 25.8%; Score 17; DB 2; Length 7;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
||
Db 6 WG 7

RESULT 14

PT0642

T-cell receptor beta chain V-D-J region (111-118) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0642
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0642
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 25.8%; Score 17; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
||
Db 5 WG 6

RESULT 15

PT0722

T-cell receptor beta chain V-D-J region (135-16) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0722
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0722
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-7 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 25.8%; Score 17; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
||
Db 6 WG 7

Search completed: August 12, 2002, 17:18:51
Job time: 329 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:20:43 ; Search time 13.13 seconds
(without alignments)
23.592 Million cell updates/sec

Title: 10-071247-2

Perfect score: 66

Sequence: 1 cywgcgyw 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 segs, 38719550 residues

Total number of hits satisfying chosen parameters: 148

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	25.8	6	1 LOK1_LOCOMI	P41491 locusta mig
2	17	25.8	8	1 AC1_THUAL	P18691 thunnus alb
3	17	25.8	8	1 LCK1_LEUMA	P21140 leucophaea
4	17	25.8	8	1 LCK2_LEUMA	P21141 leucophaea
5	17	25.8	8	1 LCK3_LEUMA	P21142 leucophaea
6	17	25.8	8	1 LCK4_LEUMA	P21143 leucophaea
7	17	25.8	8	1 LCK5_LEUMA	P19987 leucophaea
8	17	25.8	8	1 LCK6_LEUMA	P19988 leucophaea
9	17	25.8	8	1 LCK7_LEUMA	P19989 leucophaea
10	17	25.8	8	1 LCK8_LEUMA	P19990 leucophaea
11	15	22.7	4	1 OCP3_OCTMI	P58649 octopus min
12	15	22.7	5	1 UF01_MOUSE	P38639 mus musculus
13	13	19.7	8	1 ALL5_CYDPO	P82156 cydia pomon
14	12	18.2	8	1 CCKN_MACEU	P30369 macropus eu
15	11	16.7	5	1 BP77_BOTIN	P30425 bothrops in
16	11	16.7	6	1 EI01_LITRU	P82096 litoria rub
17	11	16.7	7	1 TY51_LITRU	P82065 litoria rub
18	11	16.7	7	1 WWA1_ACHFU	P35919 achatina fu
19	11	16.7	7	1 WWA2_ACHFU	P35920 achatina fu
20	11	16.7	7	1 WWA3_ACHFU	P35921 achatina fu
21	11	16.7	8	1 AKHC_GRYBI	P14086 gryllus bim
22	11	16.7	8	1 AKH_LITBAU	P25418 libellula a
23	11	16.7	8	1 AKH_MEIML	P25423 melolontha a
24	11	16.7	8	1 AKH_TABAT	P14595 tabanus atr
25	11	16.7	8	1 HTF1_PERAM	P04548 periphaneta
26	11	16.7	8	1 HTF2_PERAM	P25419 tenebrio mo
27	11	16.7	8	1 HTP_TENNO	P81707 brassica na
28	11	16.7	8	1 PLP_BRANA	P08939 pandalus bo
29	11	16.7	8	1 RPCH_PANBO	P81780 herpes simp
30	11	16.7	8	1 VGLG_HSV2B	P58648 octopus min
31	10	15.2	4	1 OCP1_OCTMI	P81864 pardachirus
32	10	15.2	5	1 PAP2_PARMA	P82073 litoria rub
33	10	15.2	5	1 RE32_LITRU	

34	10	15.2	8	1 UF06_MOUSE	P38644 mus musculus
35	9	13.6	4	1 ACH1_ACHFU	P35904 achatina fu
36	9	13.6	7	1 MNP1_LEPDE	P42984 leptinotars
37	9	13.6	8	1 ACT_CARMA	P80709 carcinus ma
38	9	13.6	8	1 GLUR_HUMAN	P02729 homo sapien
39	9	13.6	8	1 ORMY_ORCLI	P82455 orconectes
40	9	13.6	8	1 UPAA_HUMAN	P30096 homo sapien
41	8	12.1	3	1 GRWA_HUMAN	P01157 homo sapien
42	8	12.1	4	1 DCML_PSECH	P19916 pseudomonas
43	8	12.1	7	1 FAR5_HIRME	P42564 hirudo medi
44	8	12.1	7	1 UN06_PINPS	P81675 pinus pinas
45	8	12.1	8	1 AL16_CARMA	P81819 carcinus ma

ALIGNMENTS

RESULT	1
LOK1_LOCOMI	
AC	105224 segs, 38719550 residues
P41491	STANDARD; PRT; 6 AA.
DT	01-NOV-1995 (Rel. 32, Created)
DT	01-NOV-1995 (Rel. 32, Last sequence update)
DT	15-JUL-1998 (Rel. 36, Last annotation update)
DE	Locustakinin I.
OS	Locusta migratoria (Migratory locust).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC	Acridoidea; Acrididae; Acrididae; Locusta.
OX	NCBI_TaxID=7004;
RN	[1]
RP	SEQUENCE.
RC	TISSUE=Corpora cardiaca;
RX	MEDLINE=92262851; PubMed=1585017;
RA	Schoofs L., Holman G.M., Proost P., van Damme J., Hayes T.K.,
RA	de Loof A.;
RT	"Locustakinin, a novel myotropic peptide from Locusta migratoria,
RT	isolation, primary structure and synthesis.";
RL	Regul. Pept. 37:49-57(1992).
CC	!- FUNCTION: MYOTROPIC PEPTIDE. MAY BE IMPORTANT IN THE STIMULATION
CC	OF ION TRANSPORT AND INHIBITION OF DIURETIC ACTIVITY IN MALPIGHIAN
CC	TUBULES.
CC	PIR; A61068; A61068.
DR	Neuropeptide; Amidation.
KW	MOD_RES 6
FT	SEQUENCE 6 AA; 654 MW; 686365A5B9CDB000 CRC64;
SQ	AMIDATION
Query Match	25.8%; Score 17; DB 1; Length 6;
Best Local Similarity	100.0%; Pred. No. 1e+05;
Matches	2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	3 wg 4
Db	5 WG 6
RESULT	2
AC1_THUAL	
ID	AC1_THUAL
AC	105224 segs, 38719550 residues
P18691	STANDARD; PRT; 8 AA.
DT	01-NOV-1990 (Rel. 16, Created)
DT	01-NOV-1990 (Rel. 16, Last sequence update)
DT	01-NOV-1990 (Rel. 16, Last annotation update)
DE	Angiotensin-converting enzyme inhibitor.
OS	Thunnus albacares (Yellowfin tuna) (Neothunnus macropterus).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC	Acanthopterygii; Acanthopterygii; Perciformes; Scombroidei;
OC	Scombridae; Thunnus.
OX	NCBI_TaxID=8236;
RN	[1]
RP	SEQUENCE.

```
RC TISSUE=Muscle;
RX MEDLINE=88326322; PubMed=3415688;
RA Kohama Y., Matsumoto S., Oka H., Teramoto T., Okabe M., Mimura T.;
RT "Isolation of angiotensin-converting enzyme inhibitor from tuna
RL muscle.";
RL Biochem. Biophys. Res. Commun. 155:332-337(1988).
DR PIR; A31570; A31570.
SQ SEQUENCE 8 AA; 953 MW; 6AA863733051F1B7 CRC64;

Query Match 25.8%; Score 17; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 wg 4
Db 6 WG 7

RESULT 3
LCK1_LEUMA
ID LCK1_LEUMA STANDARD; PRT; 8 AA.
AC P21140;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-MAY-1991 (Rel. 18, Last annotation update)
DE Leucokinin I (L-I).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Head;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Isolation, primary structure and synthesis of two neuropeptides
RT from Leucophaea maderae: members of a new family of
RL Cephalomyotropins.";
RL Comp. Biochem. Physiol. 84C:205-211(1986).
CC -1- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC -1- ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
CC -1- SIMILARITY: TO THE OTHER LEUCOKININS.
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 893 MW; DC6365B449CDC76A CRC64;

Query Match 25.8%; Score 17; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 wg 4
Db 7 WG 8

RESULT 4
LCK2_LEUMA
ID LCK2_LEUMA STANDARD; PRT; 8 AA.
AC P21141;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-MAY-1991 (Rel. 18, Last annotation update)
DE Leucokinin II (L-II).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Head;
```

```
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Isolation, primary structure and synthesis of two neuropeptides
RT from Leucophaea maderae: members of a new family of
RL Cephalomyotropins.";
RL Comp. Biochem. Physiol. 84C:205-211(1986).
CC -1- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC -1- ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
CC -1- SIMILARITY: TO THE OTHER LEUCOKININS.
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 852 MW; DC6365A5B9C8676A CRC64;

Query Match 25.8%; Score 17; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 wg 4
Db 7 WG 8

RESULT 5
LCK3_LEUMA
ID LCK3_LEUMA STANDARD; PRT; 8 AA.
AC P21142;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-MAY-1991 (Rel. 18, Last annotation update)
DE Leucokinin III (L-III).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Head;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Primary structure and synthesis of two additional neuropeptides
RT from Leucophaea maderae: members of a new family of
RL Cephalomyotropins.";
RL Comp. Biochem. Physiol. 84C:271-276(1986).
CC -1- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC -1- ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
CC -1- SIMILARITY: TO THE OTHER LEUCOKININS.
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 910 MW; DC6365B449C866DA CRC64;

Query Match 25.8%; Score 17; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 wg 4
Db 7 WG 8

RESULT 6
LCK4_LEUMA
ID LCK4_LEUMA STANDARD; PRT; 8 AA.
AC P21143;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-MAY-1991 (Rel. 18, Last annotation update)
DE Leucokinin IV (L-IV).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
```


RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RC TISSUE-Head;
 RA Holman G.M., Cook B.J., Nachman R.J.;
 RT "Primary structure and synthesis of two additional neuropeptides
 from Leucophaea maderae: members of a new family of
 Cephalomyotropins.";
 RL Comp. Biochem. Physiol. 84C:271-276(1986).
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
 ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.
 KW Neuropeptide; Amidation.
 FT MOD_RES 8 8
 SQ SEQUENCE 8 AA; 906 MW; DC6365B1E9D5BDDA CRC64;

Query Match 25.8%; Score 17; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
 Db 7 WG 8

RESULT 7
 LCK5_LEUMA
 ID LCK5_LEUMA STANDARD; PRT; 8 AA.
 AC P19987;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Leucokinin V (L-V).
 OS Leucophaea maderae (Madeira cockroach).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
 OC Blaberoidea; Blaberidae; Leucophaea.
 OX NCBI_TaxID=6988;
 RN [1]
 RP SEQUENCE.
 RC TISSUE-Head;
 RX MEDLINE=87052651; PubMed=2877794;
 RA Holman G.M., Cook B.J., Nachman R.J.;
 RT "Isolation, primary structure, and synthesis of leucokinin V and VI:
 myotropic peptides of Leucophaea maderae.";
 RL Comp. Biochem. Physiol. 88C:27-30(1987).
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
 ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.
 DR PIR; JS0315; JS0315.
 KW Neuropeptide; Amidation.
 FT MOD_RES 8 8
 SQ SEQUENCE 8 AA; 784 MW; 736365A5B9C865B8 CRC64;

Query Match 25.8%; Score 17; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
 Db 7 WG 8

RESULT 8
 LCK6_LEUMA
 ID LCK6_LEUMA STANDARD; PRT; 8 AA.
 AC P19988;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Leucokinin VI (L-VI).
 OS Leucophaea maderae (Madeira cockroach).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
 OC Blaberoidea; Blaberidae; Leucophaea.
 OX NCBI_TaxID=6988;
 RN [1]
 RP SEQUENCE.
 RC TISSUE-Head;
 RX MEDLINE=87052651; PubMed=2877794;
 RA Holman G.M., Cook B.J., Nachman R.J.;
 RT "Isolation, primary structure, and synthesis of leucokinin V and VI:
 myotropic peptides of Leucophaea maderae.";
 RL Comp. Biochem. Physiol. 88C:27-30(1987).
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
 ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS, AND TO MANDUCA SEXTA AND
 CC HELIOTHIS ZEA ADIPOKINETIC HORMONE.
 DR PIR; JS0316; JS0316.
 KW Neuropeptide; Amidation.
 FT MOD_RES 1 1
 FT MOD_RES 8 8
 SQ SEQUENCE 8 AA; 935 MW; 9D6365B1E9D5A5A6 CRC64;

Query Match 25.8%; Score 17; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
 Db 7 WG 8

RESULT 9
 LCK7_LEUMA
 ID LCK7_LEUMA STANDARD; PRT; 8 AA.
 AC P19989;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE Leucokinin VII (L-VII).
 OS Leucophaea maderae (Madeira cockroach).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
 OC Blaberoidea; Blaberidae; Leucophaea.
 OX NCBI_TaxID=6988;
 RN [1]
 RP SEQUENCE.
 RC TISSUE-Head;
 RA Holman G.M., Cook B.J., Nachman R.J.;
 RT "Isolation, primary structure and synthesis of leucokinin VII and
 VIII: the final members of this new family of cephalomyotropic
 peptides isolated from head extracts of Leucophaea maderae.";
 RL Comp. Biochem. Physiol. 88C:31-34(1987).
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
 ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.
 DR PIR; JS0317; JS0317.
 KW Neuropeptide; Amidation.
 FT MOD_RES 8 8
 SQ SEQUENCE 8 AA; 866 MW; DC6365A5B9C8DC76A CRC64;

Query Match 25.8%; Score 17; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
 Db 7 WG 8

RESULT 10
 LCK8_LEUMA

```
ID LCK8_LEUMA STANDARD; PRT; 8 AA.
AC PI9990;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Leucokinin VIII (L-VIII);
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE.
RC TISSUE=Head;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Isolation, primary structure and synthesis of leucokinin VII and
RT VIII: the final members of this new family of cephalomyotropic
RT peptides isolated from head extracts of Leucophaea maderae.";
RL Comp. Biochem. Physiol. 88C:31-34(1987).
CC -1- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).
CC -1- SIMILARITY: TO THE OTHER LEUCOKININS.
DR PIR; JS0318.
KW Neuropeptide; Amidation.
FT MOD_RES 8 8
SQ SEQUENCE 8 AA; 902 MW; 736365AB59CAADD8 CRC64;

Query Match 25.8%; Score 17; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05; 0; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 0;

QY 3 wg 4
   ||
DB 7 WG 8

RESULT 11
OCP3_OCTMI STANDARD; PRT; 4 AA.
AC P58649;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cardioactive peptides Ocp-3/Ocp-4.
OS Octopus minor (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
OC Incirrata; Octopodidae; Octopus.
OX NCBI_TaxID=89766;
RN [1]
RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
RC TISSUE=Brain;
RA Iwakoshi E., Hisada M., Minakata H.;
RT "Cardioactive peptides isolated from the brain of a Japanese octopus,
RT Octopus minor.";
RL Peptides 21:623-630(2000).
CC -1- FUNCTION: Cardioactive; has both positive chronotropic and
CC inotropic effects on the heart. Ocp-4 is a 1000 time less
CC active than Ocp-3.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- PTM: Ocp-4 has D-Ser instead of L-Ser.
CC -1- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.
KW Hormone; D-amino acid.
FT MOD_RES 2 2
SQ SEQUENCE 4 AA; 463 MW; 6AB365B810000000 CRC64;

Query Match 22.7%; Score 15; DB 1; Length 4;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 gyw 8
   ||
```

```
DB 1 GSW 3

RESULT 12
UF01_MOUSE STANDARD; PRT; 5 AA.
ID UF01_MOUSE
AC P38639;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Unknown protein from 2D-page of fibroblasts (P19) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC TISSUE=Fibroblast;
RX MEDLINE=95009907; PubMed=7523108;
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT "Separation and sequencing of familial and novel murine proteins
RT using preparative two-dimensional gel electrophoresis.";
RL Electrophoresis 15:735-745(1994).
CC -1- MISCELLANEOUS: ON THE 2D-GEL, THE DETERMINED PI OF THIS UNKNOWN
CC NON_TER 5 5
FT SEQUENCE 5 AA; 717 MW; 7364087043100000 CRC64;

Query Match 22.7%; Score 15; DB 1; Length 5;
Best Local Similarity 66.7%; Pred. No. 1e+05; 1; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 1;

QY 6 gyw 8
   ||
DB 3 GRW 5

RESULT 13
ALL5_CYDPO STANDARD; PRT; 8 AA.
ID ALL5_CYDPO
AC P82156;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiastatin 5.
OS Cydia pomonella (Codling moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricoidea; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82600;
RN [1]
RP SEQUENCE.
RC TISSUE=Larva;
RX MEDLINE=98054539; PubMed=9392829;
RA Duve H., Johnsen A.H., Maestri J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8 8
SQ SEQUENCE 8 AA; 898 MW; 922879CABB58640D CRC64;

Query Match 19.7%; Score 13; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 gy 7
   ||
DB 3 GY 4
```

RESULT 14
CCKN_MACEU STANDARD; PRT; 8 AA.
AC P30369;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Cholecystokinin (CCK).
GN CCK.
OS Macropus eugenii (Tamar wallaby), and
OS Dasyurus viverrinus (Southeastern quoll).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9315, 9279;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=88234141; PubMed=33751140;
RA Fan Z.W., Eng J., Shaw G., Yalow R.S.;
RT "Cholecystokinin octapeptide purified from brains of Australian
RT marsupials.";
RL Peptides 9:429-431(1988).
CC -1- FUNCTION: THIS PEPTIDE HORMONE INDUCES GALL BLADDER CONTRACTION
CC AND THE RELEASE OF PANCREATIC ENZYMES IN THE GUT. ITS FUNCTION
CC IN THE BRAIN IS NOT CLEAR.
CC -1- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
DR PIR; A43001; A43001.
DR PIR; P00012; P00012.
DR InterPro; IPR001651; Gastrin.
DR PROSITE; PS00259; GASTRIN; 1.
KW Amidation; Sulfation; Hormone.
FT MOD_RES 2 2 SULFATION.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 1064 MW; DDCAA68378768B5A CRC64;

Query Match 18.2%; Score 12; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywg 4
| |
Db 2 ymg 4

RESULT 15
BPP7_BOTIN STANDARD; PRT; 5 AA.
AC P30425;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Bradykinin-potentiating peptide S5,2 (5A) (Angiotensin-converting
DE enzyme inhibitor).
OS Bothrops insularis (Island jararaca) (Queimada jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=90351557; PubMed=2386615;
RA Cintra A.C.O., Vieira C.A., Giglio J.R.;
RT "Primary structure and biological activity of bradykinin potentiating
RT peptides from Bothrops insularis snake venom.";
RL J. Protein Chem. 9:221-227(1990).
CC -1- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE
CC ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF
CC BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.

DR PIR; G37196; G37196.
KW Hypotensive agent; Venom.
FT MOD_RES 1 1
SQ SEQUENCE 5 AA; 629 MW; 776DC37326B00000 CRC64;
PYPROLIDONE CARBOXYLIC ACID.
Query Match 16.7%; Score 11; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 w 3
|
Db 3 w 3

Search completed: August 12, 2002, 17:20:43
Job time: 340 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:20:24 ; Search time 39.45 Seconds
(without alignments)
35.081 Million cell updates/sec

Title: 10-071247-2

Perfect score: 66

Sequence: 1 cywgcgyw 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 378

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_invertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	30.3	8	6 O02831	O02831_oryctolagus
2	17	25.8	7	10 O49223	O49223_glycine max
3	17	25.8	8	5 P82685	P82685_periplaneta
4	17	25.8	8	5 P82686	P82686_periplaneta
5	17	25.8	8	5 P82687	P82687_periplaneta
6	17	25.8	8	5 P82688	P82688_periplaneta
7	17	25.8	8	5 P82689	P82689_periplaneta
8	17	25.8	8	6 P82929	P82929_bos taurus
9	16	24.2	7	12 Q9YI00	Q9YI00_human adeno
10	16	24.2	7	12 Q9YI09	Q9YI09_human adeno
11	16	24.2	7	12 Q9YVE3	Q9YVE3_human adeno
12	16	24.2	8	4 Q15890	Q15890_homo sapien
13	16	24.2	8	4 Q96Q65	Q96Q65_homo sapien
14	15	22.7	7	10 Q9C5B3	Q9C5B3_arabidopsis
15	15	22.7	7	11 O55184	O55184_rattus norv
16	15	22.7	8	4 Q9Y4X6	Q9Y4X6_homo sapien

17	15	22.7	8	4 Q96RN9	Q96RN9_homo sapien
18	15	19.7	8	12 O64971	O64971_alfalifa mos
19	13	19.7	8	2 Q95R2	Q95R2_shigella dy
20	13	19.7	8	7 Q95213	Q95213_oryctolagus
21	11	16.7	6	13 P82096	P82096_litoria rub
22	11	16.7	7	8 Q95945	Q95945_saccharomyc
23	11	16.7	7	13 P82065	P82065_litoria rub
24	11	16.7	8	2 O09258	O09258_synechococc
25	11	16.7	8	2 O85406	O85406_coxiella bu
26	11	16.7	8	4 Q9BYV5	Q9BYV5_homo sapien
27	11	16.7	8	4 Q15888	Q15888_homo sapien
28	11	16.7	8	5 Q9VRD2	Q9VRD2_drosophila
29	11	16.7	8	6 Q9TRY3	Q9TRY3_sus sp. ins
30	11	16.7	8	6 Q9BFC3	Q9BFC3_didelphis m
31	11	16.7	8	6 Q9BFC2	Q9BFC2_macropus eu
32	11	16.7	8	6 Q9BFC1	Q9BFC1_choloepus h
33	11	16.7	8	6 Q9BFC0	Q9BFC0_choloepus d
34	11	16.7	8	6 Q9BFB9	Q9BFB9_euphractus
35	11	16.7	8	6 Q9BFB8	Q9BFB8_chaetophrac
36	11	16.7	8	6 Q9BFB7	Q9BFB7_tamandua te
37	11	16.7	8	6 Q9BFB6	Q9BFB6_myrmecophag
38	11	16.7	8	6 Q9BFB5	Q9BFB5_einaceus c
39	11	16.7	8	6 Q9BFB4	Q9BFB4_taipa alta
40	11	16.7	8	6 Q9BFB3	Q9BFB3_condylura c
41	11	16.7	8	6 Q9BFB2	Q9BFB2_sorex arane
42	11	16.7	8	6 Q9BFB1	Q9BFB1_echinops te
43	11	16.7	8	6 Q9BFB0	Q9BFB0_trichechus
44	11	16.7	8	6 Q9BFA9	Q9BFA9_procavia ca
45	11	16.7	8	6 Q9BFA8	Q9BFA8_loxodonta a

ALIGNMENTS

RESULT 1

O02831 PRELIMINARY; PRT; 8 AA.
ID O02831
AC O02831;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PRO ALPHA 1 TYPE III COLLAGEN PROTEIN (FRAGMENT).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96377339; PubMed=8783186;
RA Metsaranta M., Kujala U.M., Pelliniemi L., Osterman H., Aho H.,
RA Vuorio E.;
RT "Evidence for insufficient chondrocytic differentiation during repair
of full-thickness defects of articular cartilage.";
RL Matrix Biol. 15:39-47(1996).
DR EMBL; S83371; AAD14433.1; -.
KW Collagen.
FT NON_TER
SQ SEQUENCE 8 AA; 1028 MW; B859C7272EA77371 CRC64;

Query Match 30.3%; Score 20; DB 6; Length 8;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywgc 5
: | |
Db 1 HWPC 4

RESULT 2

O49223 PRELIMINARY; PRT; 7 AA.
ID O49223
AC O49223;

```
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HMG-1-LIKE PROTEIN (FRAGMENT).
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ESSEX; TISSUE=ROOT;
RX MEDLINE=91367679; PubMed=1891369;
RA Laux T., Goldberg R.B.;
RT "A plant DNA binding protein shares highly conserved sequence motifs
with HMG-box proteins.";
RL Nucleic Acids Res. 19:4769-4769(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ESSEX; TISSUE=ROOT;
RA Mahalingam R., Knap H.T.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF047050; AAC03556.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 850 MW; 6AAAAAB378637810 CRC64;

Query Match      25.8%; Score 17; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
DB 2 WG 3

RESULT 3
ID P82685 PRELIMINARY; PRT; 8 AA.
AC P82685;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE KININ-1 (PEA-K-1).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattoidea; Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=CORPORA CARDIACA;
RX MEDLINE=98010462; PubMed=9350979;
RA Predel R., Kellner R., Rapus J., Penzlin H., Gaede G.;
RT "Isolation and structural elucidation of eight kinins from the
retrocerebral complex of the American cockroach, Periplaneta
americana.";
RL Regul. Pept. 71:199-205(1997).
CC -1- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
(MYOTROPIC ACTIVITY).
CC -1- SIMILARITY: BELONGS TO THE KININ FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 950 MW; 326365B449D5A774 CRC64;

Query Match      25.8%; Score 17; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
DB 7 WG 8
```

```
RESULT 4
ID P82686 PRELIMINARY; PRT; 8 AA.
AC P82686;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE KININ-2 (PEA-K-2).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattoidea; Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=CORPORA CARDIACA;
RX MEDLINE=98010462; PubMed=9350979;
RA Predel R., Kellner R., Rapus J., Penzlin H., Gaede G.;
RT "Isolation and structural elucidation of eight kinins from the
retrocerebral complex of the American cockroach, Periplaneta
americana.";
RL Regul. Pept. 71:199-205(1997).
CC -1- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
(MYOTROPIC ACTIVITY).
CC -1- SIMILARITY: BELONGS TO THE KININ FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 856 MW; DC6365A5B9D5BDDA CRC64;

Query Match      25.8%; Score 17; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 wg 4
DB 7 WG 8

RESULT 5
ID P82687 PRELIMINARY; PRT; 8 AA.
AC P82687;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE KININ-3 (PEA-K-3).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattoidea; Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=CORPORA CARDIACA;
RX MEDLINE=98010462; PubMed=9350979;
RA Predel R., Kellner R., Rapus J., Penzlin H., Gaede G.;
RT "Isolation and structural elucidation of eight kinins from the
retrocerebral complex of the American cockroach, Periplaneta
americana.";
RL Regul. Pept. 71:199-205(1997).
CC -1- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
(MYOTROPIC ACTIVITY).
CC -1- SIMILARITY: BELONGS TO THE KININ FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 909 MW; DC6365B449D5A76A CRC64;

Query Match      25.8%; Score 17; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
```

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 wg 4
||

Db 7 WG 8

RESULT 6

ID P82688 PRELIMINARY; PRT; 8 AA.
AC P82688;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE KININ-4 (PEA-K-4).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattelloidea; Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
[1]
RN SEQUENCE, AND FUNCTION.
RP TISSUE=CORPORA CARDIACA;
RC MEDLINE=98010462; PubMed=9350979;
RX Predel R., Kellner R., Rapus J., Penzlin H., Gaede G.;
RT "Isolation and structural elucidation of eight kinins from the
retrocerebral complex of the American cockroach, Periplaneta
americana.";
RL Regul. Pept. 71:199-205(1997).
CC -1- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
(MYOTROPIC ACTIVITY).
CC -1- SIMILARITY: BELONGS TO THE KININ FAMILY.
CC Neuropeptide; Amidation.
KW MOD_RES 8 8 AMIDATION.
FT SEQUENCE 8 AA; 839 MW; 736365A5B9D6DDDD8 CRC64;

Query Match 25.8%; Score 17; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 wg 4
||

Db 7 WG 8

RESULT 7

ID P82689 PRELIMINARY; PRT; 8 AA.
AC P82689;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE KININ-5 (PEA-K-5).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattelloidea; Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
[1]
RN SEQUENCE, AND FUNCTION.
RP TISSUE=CORPORA CARDIACA;
RC MEDLINE=98010462; PubMed=9350979;
RX Predel R., Kellner R., Rapus J., Penzlin H., Gaede G.;
RT "Isolation and structural elucidation of eight kinins from the
retrocerebral complex of the American cockroach, Periplaneta
americana.";
RL Regul. Pept. 71:199-205(1997).
CC -1- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
(MYOTROPIC ACTIVITY).
CC -1- SIMILARITY: BELONGS TO THE KININ FAMILY.
CC Neuropeptide; Amidation.
KW MOD_RES 8 8 AMIDATION.

SQ SEQUENCE 8 AA; 865 MW; C76365B449CDC775 CRC64;

Query Match 25.8%; Score 17; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 wg 4
||

Db 7 WG 8

RESULT 8

ID P82929 PRELIMINARY; PRT; 8 AA.
AC P82929;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE MITOCHONDRIAL 28S RIBOSOMAL PROTEIN S34 (MRP-S34) (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=LIVER;
RA Koc E.C., Burkhardt W., Blackburn K., Moseley A., Spremulli L.L.;
RT "Small subunit of the mammalian mitochondrial ribosome. Identification
of the full complement ribosomal proteins present.";
RL Submitted (DEC-2000) to the SWISS-PROT data bank.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
KW Ribosomal protein; Mitochondrion.
FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 935 MW; 9639D1A72058637D CRC64;

Query Match 25.8%; Score 17; DB 6; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 wg 4
||

Db 2 WG 3

RESULT 9

ID Q9YIRO PRELIMINARY; PRT; 7 AA.
AC Q9YIRO;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE PVI CORE PROTEIN (FRAGMENT).
GN PVI.
OS Human adenovirus type 7a.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=85755;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=KN T96-0620, S-1058, AND CL 68578;
RA Crawford-Miksz L.K., Nang R.N., Schnurr D.P.;
RT "Molecular surveillance of strain variation in adenoviruses causing
acute respiratory disease, AV 4 and AV 7a.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065068; AAD03668.1; -.
DR EMBL; AF065066; AAD03664.1; -.
DR EMBL; AF065067; AAD03666.1; -.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 980 MW; 7B5EA414140322A0 CRC64;

Query Match 24.2%; Score 16; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CY 2
||
Db 6 CY 7

RESULT 10
Q9YIQ9 PRELIMINARY; PRT; 7 AA.
AC Q9YIQ9;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE PVI CORE PROTEIN (FRAGMENT).
GN PVI.
OS Human adenovirus type 4.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=28280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Z-G 95-873, RI-67, AND 55142;
RA Crawford-Miksza L.K., Nang R.N., Schnurr D.P.;
RT "Molecular surveillance of strain variation in adenoviruses causing acute respiratory disease, AV 4 and AV 7a.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065064; AAD03659.1; -.
DR EMBL; AF065062; AAD03653.1; -.
DR EMBL; AF065063; AAD03656.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 980 MW; 7B5EA414140322A0 CRC64;

Query Match 24.2%; Score 16; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CY 2
||
Db 6 CY 7

RESULT 11
Q9YVE3 PRELIMINARY; PRT; 7 AA.
AC Q9YVE3;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PVI CORE PROTEIN (FRAGMENT).
GN PVI.
OS Human adenovirus type 7.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GOMEN;
RX MEDLINE=99175282; PubMed=10074533;
RA Crawford-Miksza L.K., Nang R.N., Schnurr D.P.;
RT "Strain variation in adenovirus serotypes 4 and 7a causing acute respiratory disease.";
RL J. Clin. Microbiol. 37:1107-1112(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-GOMEN;
RA Crawford-Miksza L.K.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065065; AAD03662.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 980 MW; 7B5EA414140322A0 CRC64;

Query Match 24.2%; Score 16; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CY 2
||
Db 6 CY 7

RESULT 12
Q15890 PRELIMINARY; PRT; 8 AA.
AC Q15890;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE (CLONE XP19G12A) (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M., Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y., Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L32083; AAA73880.1; -.
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 975 MW; 605EA6C5BEA5A2D3 CRC64;

Query Match 24.2%; Score 16; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CY 2
||
Db 7 CY 8

RESULT 13
Q96Q65 PRELIMINARY; PRT; 8 AA.
AC Q96Q65;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MITOCHONDRIAL RIBOSOMAL PROTEIN L39 (FRAGMENT).
GN MRPL39.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429115; PubMed=11543634;
RA Kenmochi N., Suzuki T., Uechi T., Magoori M., Kuniba M., Higa S., Watanabe K., Tanaka T.;
RT "The human mitochondrial ribosomal protein genes: Mapping of 54 genes to the chromosomes and implications for human disorders.";
RL Genomics 77:65-70(2001).
DR EMBL; AB051346; BAB54936.1; -.
RW Ribosomal protein.
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 875 MW; 47BB5EB9CDD865B0 CRC64;

• • •

Query Match 24.2%; Score 16; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CY 2
||
DB 6 CY 7

RESULT 14
Q9C5B3
ID Q9C5B3 PRELIMINARY; PRT; 7 AA.
AC Q9C5B3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 0.7 KDA PROTEIN (FRAGMENT).
GN DIDI 10A-2B.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT;
RX MEDLINE=21171025; PubMed=11277426;
RA Vercauteren I., Van Der Schueren E., Van Montagu M., Gheysen G.,
RT "Arabidopsis thaliana genes expressed in the early compatible
interaction with root-knot nematodes.";
RL Mol. Plant Microbe Interact. 14:288-299(2001).
DR EMBL: AJ286350; CAB71014.2; -.
KW Hypothetical protein.
FT NON_TER 1 1
FT NON_TER 7 7
SQ SEQUENCE 7 AA; 719 MW; 6732C7287EB325D0 CRC64;

Query Match 22.7%; Score 15; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CG 6
||
DB 3 CG 4

RESULT 15
O55184
ID O55184 PRELIMINARY; PRT; 7 AA.
AC O55184;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORPHAN RECEPTOR TR4-NS (FRAGMENT).
GN TR4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=96198747; PubMed=8612486;
RA Yoshikawa T., Makino S., Gao X.M., Xing G.Q., Chuang D.M.,
RA Detera-Wadleigh S.D.;
RT "Splice variants of rat TR4 orphan receptor: differential expression
of novel sequences in the 5'-untranslated region and C-terminal
domain.";
RL Endocrinology 137:1562-1571(1996).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=96299786; PubMed=8661150;
RA Yoshikawa T., DuPont B.R., Leach R.J., Detera-Wadleigh S.D.;
RT "New variants of the human and rat nuclear hormone receptor, TR4:
expression and chromosomal localization of the human gene.";
RL Genomics 35:361-366(1996).
DR EMBL: U59454; AAB91433.1; -.
KW Receptor.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 663 MW; 6DDAA8787EB05350 CRC64;

Query Match 22.7%; Score 15; DB 11; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CG 6
||
DB 3 CG 4

Search completed: August 12, 2002, 17:20:24
Job time: 346 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:14:32 ; Search time 48.76 Seconds
(without alignments)
18.224 Million cell updates/sec

Title: 10-071247-2

Perfect score: 66

Sequence: 1 cywgcgyw 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 69368

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	100.0	8	20	AA1980
2	66	100.0	8	21	AA1981
3	45	68.2	8	20	AA1982
4	45	68.2	8	20	AA1983
5	42	63.6	7	20	AA1984
6	40	60.6	7	20	AA1985
7	39	59.1	8	22	AA1986
8	39	59.1	8	22	AA1987
9	37	56.1	8	21	AA1988
10	35	53.0	8	22	AA1989
11	34	51.5	6	21	AA1990

12	34	51.5	7	13	AA1980	Somatostatin analog
13	34	51.5	7	15	AA1981	Somatostatin analog
14	34	51.5	7	18	AA1982	Somatostatin analog
15	34	51.5	7	19	AA1983	Somatostatin analog
16	34	51.5	7	19	AA1984	Somatostatin analog
17	34	51.5	7	19	AA1985	Somatostatin analog
18	34	51.5	7	19	AA1986	Somatostatin analog
19	34	51.5	7	19	AA1987	Somatostatin analog
20	34	51.5	7	20	AA1988	Somatostatin analog
21	34	51.5	7	20	AA1989	Somatostatin analog
22	34	51.5	7	20	AA1990	Somatostatin analog
23	34	51.5	7	20	AA1991	Somatostatin analog
24	34	51.5	7	20	AA1992	Somatostatin analog
25	34	51.5	7	20	AA1993	Somatostatin analog
26	34	51.5	7	20	AA1994	Somatostatin analog
27	34	51.5	7	20	AA1995	Somatostatin analog
28	34	51.5	7	20	AA1996	Somatostatin analog
29	34	51.5	7	20	AA1997	Somatostatin analog
30	34	51.5	7	20	AA1998	Somatostatin analog
31	34	51.5	7	21	AA1999	Somatostatin analog
32	34	51.5	7	21	AA2000	Somatostatin analog
33	34	51.5	7	21	AA2001	Somatostatin analog
34	34	51.5	7	21	AA2002	Somatostatin analog
35	34	51.5	7	21	AA2003	Somatostatin analog
36	34	51.5	7	21	AA2004	Somatostatin analog
37	34	51.5	7	21	AA2005	Somatostatin analog
38	34	51.5	7	22	AA2006	Somatostatin analog
39	34	51.5	7	22	AA2007	Somatostatin analog
40	34	51.5	7	22	AA2008	Somatostatin analog
41	34	51.5	8	22	AA2009	Somatostatin analog
42	34	51.5	8	22	AA2010	Somatostatin analog
43	33	50.0	5	22	AA2011	Somatostatin analog
44	33	50.0	8	22	AA2012	Somatostatin analog
45	32	48.5	8	22	AA2013	Somatostatin analog

ALIGNMENTS

RESULT 1		
AAAY03715		
ID	AAAY03715	standard; peptide; 8 AA.
XX		
AC	AAAY03715;	
XX		
DT	08-JUN-1999	(first entry)
XX		
DE	Fluorine-18 (F-18)	labeled peptide 2.
XX		
KW	18F radionuclide;	targeting vector; positron emission tomography; F-18;
KW	radiolabeling;	thiol; fluorine-18.
XX		
OS	Synthetic.	
XX		
Key		Location/Qualifiers
FH	Misc-difference 1	
FT		/note= "N-terminal acetylation; optionally has a free
FT		or protected thiol group"
FT	Misc-difference 2	
FT		/note= "D-form residue"
FT	Misc-difference 3	
FT		/note= "D-form residue"
FT	Misc-difference 5	
FT		/note= "D-form residue; optionally has a free or
FT		protected thiol group"
FT	Misc-difference 7	
FT		/note= "D-form residue"
FT	Misc-difference 8	
FT		/note= "D-form residue"
XX		
XX	WO9911590-A1.	
PN		
XX		
PD	11-MAR-1999.	

```
XX 03-SEP-1998; 98WO-US18268.
XX
XX 03-SEP-1997; 97US-0057485.
XX
XX (IMMU-) IMMUNOMEDICS INC.
XX
XX Griffiths GL;
XX
XX WPI; 1999-228967/19.
XX
XX Radiolabeling thiol-containing peptides with fluorine-18
XX
XX Claim 14; Page 15; 22pp; English.
XX
XX The invention relates to a method for incorporating 18F radionuclide
XX into peptide-containing targeting vectors for use in clinical positron
XX emission tomography. Radiolabeling thiol-containing peptides with
XX fluorine-18 (F-18) comprises reacting a peptide comprising a free thiol
XX group with a labeling reagent of formula: 18F-(CH2)m-CR1R2-(CH2)n-X, or
XX a fluorinated alkene in which at least one of the two double bonded
XX carbon atoms bears at least one leaving group comprising I, Br, Cl,
XX azide, tosylate, mesylate, nosylate or triflate. n, m = 0-2; n+m = 0-2;
XX X = I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, maleimide
XX (optionally substituted by 1-2 alkyl) or 3-sulfomaleimide; R1, R2 = I,
XX Br, Cl, azide, tosylate, mesylate, nosylate, triflate, H, CONH2, COOH,
XX OH, sulfonic acid, tertiary amine, quaternary ammonium, alkyl
XX (optionally substituted by CONH2, COOH, OH, sulfonic acid, tertiary amine
XX or quaternary ammonium), COOR', CONR'2 or COR'; and R' = 1-6C alkyl or
XX phenyl. The method is used for Radiolabeling peptide-containing
XX targeting vectors such as proteins, antibodies, antibody fragments and
XX receptor- targeted peptides for use in routine clinical positron
XX emission tomography. The method is simple and efficient. The method uses
XX the unique property of the free thiol groups which are rapidly alkylated
XX at neutral pH and moderate temperature. Sequences AAY03714-716 represent
XX examples of F-18 labeled peptides used in the method of detecting a
XX tissue.
XX
XX SQ Sequence 8 AA;
XX
XX Query Match 100.0%; Score 66; DB 20; Length 8;
XX Best Local Similarity 100.0%; Pred. No. 6.4e+05;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 cywgcgyw 8
XX ||||||
XX Db 1 cywgcgyw 8
XX
XX RESULT 2
XX ID AAY76817 standard; peptide; 8 AA.
XX
XX AC AAY76817;
XX
XX DT 28-APR-2000 (first entry)
XX
XX Immunogenic peptide for bi-specific antibody recognition.
XX
XX Immunogenic peptide; bi-specific antibody; diagnosis; immune response;
XX diseased tissue identification; therapy.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 1
XX \ /note= "acetylated; modified with free amino acid group,
XX protected amino acid group, chelating agent or a
XX metal-chelate complex"
XX Misc-difference 2
XX /note= "D-form residue"
XX Misc-difference 3
XX
XX FT Misc-difference 5
XX /note= "D-form residue; modified with free amino acid
XX group, protected amino acid group, chelating
XX agent or a metal-chelate complex"
XX
XX FT Misc-difference 7
XX /note= "D-form residue"
XX
XX FT Misc-difference 8
XX /note= "D-form residue"
XX
XX PN WO9966951-A2.
XX
XX XX 29-DEC-1999.
XX
XX XX 22-JUN-1999; 99WO-US13879.
XX
XX PR 22-JUN-1998; 98US-0090142.
XX
XX PR 14-OCT-1998; 98US-0104156.
XX
XX PA (IMMU-) IMMUNOMEDICS INC.
XX
XX PI Hansen HJ, Griffiths GL, Leung S, McBride WJ, Qu Z;
XX
XX WPI; 2000-160561/14.
XX
XX Bi-specific antibodies that bind specific target tissue and targeted
XX conjugates -
XX
XX Claim 22; Page 61; 76pp; English.
XX
XX This sequence represents an immunogenic peptide for a bi-specific
XX antibody. The invention relates to a method of treating or identifying
XX diseased tissues in a patient comprising administering a bi-specific
XX antibody (or fragment) having at least 1 arm (A) that specifically
XX binds a targeted tissue and at least 1 arm (B) that specifically binds a
XX targetable conjugate. The methods and bi-specific antibodies and fusion
XX proteins are useful for pre-targeting methods of diagnosis and therapy.
XX It is advantageous to raise bi-specific antibodies against a targetable
XX conjugate that is capable of carrying at least 1 diagnostic or
XX therapeutic agent. The characteristics of the chelator, metal chelate
XX complex, therapeutic agent or diagnostic agent can be varied to
XX accommodate differing applications without raising new bi-specific
XX antibodies for each new application. The targetable conjugate is selected
XX to elicit sufficient immune responses and also for rapid in vivo
XX clearance when used within the bi-specific antibody targeting method.
XX
XX SQ Sequence 8 AA;
XX
XX Query Match 100.0%; Score 66; DB 21; Length 8;
XX Best Local Similarity 100.0%; Pred. No. 6.4e+05;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 cywgcgyw 8
XX ||||||
XX Db 1 cywgcgyw 8
XX
XX RESULT 3
XX ID AAY03714 standard; peptide; 8 AA.
XX
XX AC AAY03714;
XX
XX DT 08-JUN-1999 (first entry)
XX
XX DE Fluorine-18 (F-18) labeled peptide 1.
XX
XX KW 18F radionuclide; targeting vector; positron emission tomography; F-18;
XX radiolabeling; thiol; fluorine-18.
XX
XX OS Synthetic.
XX
```

FH Key Location/Qualifiers
 FT Misc-difference 1 /note= "optionally has a free or protected amino acid group"
 FT Misc-difference 2 /note= "D-form residue"
 FT Misc-difference 3 /note= "D-form residue"
 FT Misc-difference 5 /note= "D-form residue"
 FT Misc-difference 7 /note= "D-form residue; optionally has a free or protected amino acid group"
 FT Misc-difference 8 /note= "D-form residue"
 FT Misc-difference 8 /note= "D-form residue"
 XX WO9911590-A1.
 XX 11-MAR-1999.
 PD 03-SEP-1998; 98WO-US18268.
 XX 03-SEP-1997; 97US-0057485.
 XX (IMMU-) IMMUNOMEDICS INC.
 PA Griffiths GL;
 PI WPI; 1999-228967/19.
 XX Radiolabeling thiol-containing peptides with fluorine-18
 PS Claim 13; Page 15; 22pp; English.
 CC The invention relates to a method for incorporating 18F radionuclide into peptide-containing targeting vectors for use in clinical positron emission tomography. Radiolabeling thiol-containing peptides with fluorine-18 (F-18) comprises reacting a peptide comprising a free thiol group with a labeling reagent of formula: 18F-(CH₂)_m-CR1R2-(CH₂)_n-X, or a fluorinated alkene in which at least one of the two double bonded carbon atoms bears at least one leaving group comprising I, Br, Cl, azide, tosylate, mesylate, nosylate or triflate. n, m = 0-2; X = I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, maleimide (optionally substituted by 1-2 alkyl) or 3-sulfonamide; R1, R2 = I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, H, CONH₂, COOH, OH, sulfonic acid, tertiary amine, quaternary ammonium, alkyl (optionally substituted by CONH₂, COOH, OH, sulfonic acid, tertiary amine or quaternary ammonium), COOR', CONR'2 or COR'; and R' = 1-6C alkyl or phenyl. The method is used for Radiolabeling peptide-containing targeting vectors such as proteins, antibodies, antibody fragments and receptor-targeted peptides for use in routine clinical positron emission tomography. The method is simple and efficient. The method uses the unique property of the free thiol groups which are rapidly alkylated at neutral pH and moderate temperature. Sequences AAY03714-716 represent examples of F-18 labeled peptides used in the method of detecting a tissue.
 XX Sequence 8 AA;
 SQ
 Query Match 68.28; Score 45; DB 20; Length 8;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 ywgcyw 8
 ||| |||
 Db 2 ywgkyw 8

RESULT 4

AAY03716
 ID AAY03716 standard; peptide; 8 AA.
 XX

AC AAY03716;
 XX 08-JUN-1999 (first entry)
 XX Fluorine-18 (F-18) labeled peptide 3.
 DE 18F radionuclide; targeting vector; positron emission tomography; F-18;
 KW radiolabeling; thiol; fluorine-18.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetylation"
 FT Misc-difference 2 /note= "D-form residue; D-iodo-Tyr"
 FT Misc-difference 3 /note= "D-form residue"
 FT Misc-difference 5 /note= "D-form residue"
 FT Misc-difference 7 /note= "D-form residue; optionally acetylated"
 FT Misc-difference 8 /note= "D-form residue; D-iodo-Tyr"
 FT Misc-difference 8 /note= "D-form residue"
 XX WO9911590-A1.
 PN 11-MAR-1999.
 PD 03-SEP-1998; 98WO-US18268.
 PF 03-SEP-1997; 97US-0057485.
 XX (IMMU-) IMMUNOMEDICS INC.
 PA Griffiths GL;
 PI WPI; 1999-228967/19.
 XX Radiolabeling thiol-containing peptides with fluorine-18
 PS Claim 15; Page 15; 22pp; English.
 XX The invention relates to a method for incorporating 18F radionuclide into peptide-containing targeting vectors for use in clinical positron emission tomography. Radiolabeling thiol-containing peptides with fluorine-18 (F-18) comprises reacting a peptide comprising a free thiol group with a labeling reagent of formula: 18F-(CH₂)_m-CR1R2-(CH₂)_n-X, or a fluorinated alkene in which at least one of the two double bonded carbon atoms bears at least one leaving group comprising I, Br, Cl, azide, tosylate, mesylate, nosylate or triflate. n, m = 0-2; X = I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, maleimide (optionally substituted by 1-2 alkyl) or 3-sulfonamide; R1, R2 = I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, H, CONH₂, COOH, OH, sulfonic acid, tertiary amine, quaternary ammonium, alkyl (optionally substituted by CONH₂, COOH, OH, sulfonic acid, tertiary amine or quaternary ammonium), COOR', CONR'2 or COR'; and R' = 1-6C alkyl or phenyl. The method is used for Radiolabeling peptide-containing targeting vectors such as proteins, antibodies, antibody fragments and receptor-targeted peptides for use in routine clinical positron emission tomography. The method is simple and efficient. The method uses the unique property of the free thiol groups which are rapidly alkylated at neutral pH and moderate temperature. Sequences AAY03714-716 represent examples of F-18 labeled peptides used in the method of detecting a tissue.
 XX Sequence 8 AA;
 SQ

Query Match 68.28; Score 45; DB 20; Length 8;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 ywgcgyw 8
Db 2 ywgcgyw 8

RESULT 5

AA42908
ID AAY42908 standard; peptide: 7 AA.

XX AC
XX AAY42908;

DT 20-DEC-1999 (first entry)

XX Somatostatin derivative VZ-934 for antiinflammatory or analgesic use.

XX Somatostatin; antiinflammatory; analgesic; rhinitis; bronchitis;
KW bronchial asthma; arthritis; allergic conjunctivitis; urticaria;
KW colitis; psoriasis.

XX OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "D-form residue"

FT Misc-difference 4 /note= "D-form residue"

FT Disulfide-bond 2..6

FT Modified-site 7

FT /note= "Thr-NH2"

XX EP952159-A2.

XX 27-OCT-1999.

XX 23-APR-1999; 99EP-0107392.

XX 24-APR-1998; 98HU-0000970.

XX (BIOS-) BIOSTATIN GYOGYSZERKUTATO-PEJLESZTO KFT.

XX Keri G, Szolcsanyi J, Pinter E, Helyes Z, Erchegeyi J, Horvath A;

PI Horvath J, Teplan I, Orfi L;

XX WPI; 1999-582791/50.

XX Use of somatostatin derivatives and phenylhydrazone derivatives for

PT preparation of medicaments with neurogenic and/or non-neurogenic

PT antiinflammatory and/or analgesic effects -

XX Claim 8; Page 13; 20pp; English.

XX The invention relates to the use of certain heptapeptide somatostatin

CC derivatives and/or phenylhydrazone derivatives for the preparation

CC of medicaments possessing neurogenic and/or non-neurogenic

CC antiinflammatory and/or analgesic effects. The derivatives are used

CC for treating pain and for treating neurogenic inflammation in the

CC pathomechanism of rhinitis, bronchitis, bronchial asthma, arthritis,

CC allergic conjunctivitis, urticaria, colitis and psoriasis. The present

CC sequence is a preferred example of a heptapeptide somatostatin

CC derivative for use in the invention.

XX Sequence 7 AA;

SQ

Query Match 63.6%; Score 42; DB 20; Length 7;

Best Local Similarity 100.0%; Pred. No. 6.4e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cywgc 5

Db 2 cywgc 6

XX

AC

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

RESULT 6

AA42915
ID AAY42915 standard; peptide: 7 AA.

XX AC
XX AAY42915;

DT 20-DEC-1999 (first entry)

XX Somatostatin derivative VZ-1038 for antiinflammatory or analgesic use.

XX Somatostatin; antiinflammatory; analgesic; rhinitis; bronchitis;
KW bronchial asthma; arthritis; allergic conjunctivitis; urticaria;
KW colitis; psoriasis.

XX OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "D-form residue"

FT Misc-difference 4 /note= "D-form residue"

FT Disulfide-bond 2..6

FT Modified-site 7

FT /note= "Gly-NH2"

XX EP952159-A2.

XX 27-OCT-1999.

XX 23-APR-1999; 99EP-0107392.

XX 24-APR-1998; 98HU-0000970.

XX (BIOS-) BIOSTATIN GYOGYSZERKUTATO-PEJLESZTO KFT.

XX Keri G, Szolcsanyi J, Pinter E, Helyes Z, Erchegeyi J, Horvath A;

PI Horvath J, Teplan I, Orfi L;

XX WPI; 1999-582791/50.

XX Use of somatostatin derivatives and phenylhydrazone derivatives for

PT preparation of medicaments with neurogenic and/or non-neurogenic

PT antiinflammatory and/or analgesic effects -

XX Claim 8; Page 13; 20pp; English.

XX The invention relates to the use of certain heptapeptide somatostatin

CC derivatives and/or phenylhydrazone derivatives for the preparation

CC of medicaments possessing neurogenic and/or non-neurogenic

CC antiinflammatory and/or analgesic effects. The derivatives are used

CC for treating pain and for treating neurogenic inflammation in the

CC pathomechanism of rhinitis, bronchitis, bronchial asthma, arthritis,

CC allergic conjunctivitis, urticaria, colitis and psoriasis. The present

CC sequence is a preferred example of a heptapeptide somatostatin

CC derivative for use in the invention.

XX Sequence 7 AA;

SQ

Query Match 60.6%; Score 40; DB 20; Length 7;

Best Local Similarity 83.3%; Pred. No. 6.4e+05;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 cywgcg 6

Db 2 cywkcq 7

XX

AC

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

AA59639;

QY 1 cywgcgy 7
Db 2 cywkcty 8

RESULT 9
AA76818
ID AAY76818 standard; peptide; 8 AA.
AC AAY76818;
DT 28-APR-2000 (first entry)
DE Immunogenic peptide for bi-specific antibody recognition.
XX
XX Immunogenic peptide; bi-specific antibody; diagnosis; immune response;
KW diseased tissue identification; therapy.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "acetylated"
FT Misc-difference 2 /note= "D-form iodo-Tyr"
FT Misc-difference 3 /note= "D-form residue"
FT Misc-difference 5 /note= "acetylated D-form residue"
FT Misc-difference 7 /note= "D-form iodo-Tyr"
FT Misc-difference 8 /note= "D-form residue"
XX
XX WO9966951-A2.
PN
XX
XX 29-DEC-1999.
XX
XX 22-JUN-1999; 99WO-US13879.
XX
XX 22-JUN-1998; 98US-0090142.
XX 14-OCT-1998; 98US-0104156.
XX
XX (IMMU-) IMMUNOMEDICS INC.
XX
XX Hansen HJ, Griffiths GL, Leung S, McBride WJ, Qu Z;
XX WPI; 2000-160561/14.
XX
XX Bi-specific antibodies that bind specific target tissue and targeted
XX conjugates -
XX
XX Claim 23; Page 61; 76pp; English.
XX
XX This sequence represents an immunogenic peptide for a bi-specific
XX antibody. The invention relates to a method of treating or identifying
XX diseased tissues in a patient comprising administering a bi-specific
XX antibody (or fragment) having at least 1 arm (A) that specifically
XX binds a targeted tissue and at least 1 arm (B) that specifically binds a
XX targetable conjugate. The methods and bi-specific antibodies and fusion
XX proteins are useful for pre-targeting methods of diagnosis and therapy.
XX It is advantageous to raise bi-specific antibodies against a targetable
XX conjugate that is capable of carrying at least 1 diagnostic or
XX therapeutic agent. The characteristics of the chelator, metal chelate
XX complex, therapeutic agent or diagnostic agent can be varied to
XX accommodate differing applications without raising new bi-specific
XX antibodies for each new application. The targetable conjugate is selected
XX to elicit sufficient immune responses and also for rapid in vivo
XX clearance when used within the bi-specific antibody targeting method.
XX
XX Sequence 8 AA;

Query Match 56.1%; Score 37; DB 21; Length 8;
Best Local Similarity 71.4%; Pred. No. 6.4e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ywgcgyw 8
Db 2 ywhkgyw 8

RESULT 10
AAB59641
ID AAB59641 standard; Peptide; 8 AA.
XX
AC AAB59641;
XX
DT 23-MAR-2001 (first entry)
XX
XX Neuromedin B and somatostatin analogue #34.
XX
XX Neuromedin B; NMB; somatostatin; analogue; Helicobacter pylori; cancer;
KW cyclic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /label= OTHER
FT /note= "optionally N-(3-indolylacetyl)-L-phenylalanine,
FT N-(2-aminoethyl)-N-(2-thymine)-L-oxo-ethyl)-glycine,
FT N-(2-aminoethyl)-N-(2-cytosine)-L-oxo-ethyl)-glycine,
FT 5-(4-methyl-2-nitrophenyl)-2-furoic acid,
FT isonipecotic acid or nicotinuric acid"
FT Disulfide-bond 2..6
FT /note= "disulfide bond cyclises the peptide"
FT
FT Misc-difference 2 /note= "D-form residue"
FT Misc-difference 4 /note= "D-form residue"
FT Misc-difference 6 /note= "D-form residue"
FT Misc-difference 7 /note= "optionally D-form residue"
FT Modified-site 7 /label= OTHER
FT /note= "modified by Bzl"
FT Modified-site 8 /label= OTHER
FT /note= "C-terminal amide"
XX
XX WO200075186-A1.
XX
XX 14-DEC-2000.
XX
XX 05-JUN-2000; 2000WO-US15396.
XX
XX 04-JUN-1999; 99US-0137655.
XX
XX (BTOM-) BIOMEDICAL INC.
XX
XX Sadat-Aalae D, Morgan BA;
XX
XX WPI; 2001-146787/15.
XX
XX New polypeptide compounds are somatostatin and neuromedin B
XX agonists, for treating a wide range of disorders e.g. cancer,
XX gastrointestinal disorders and inflammatory disorders -
XX
XX Claim 15; Page 73; 85pp; English.
XX
XX The present invention provides analogues of somatostatin and neuromedin B
XX (NMB). These can be used in the treatment of various cancers, anorexia,
XX hypothyroidism, hyperaldosteronism, Helicobacter pylori proliferation,
XX acromegaly, restenosis, Crohn's disease, systemic sclerosis, external and
XX internal pancreatic pseudocysts and ascites, nesidoblastosis,

CC hyperinsulinism, Zollinger-Ellison syndrome, diarrhoea, scleroderma,
 CC irritable bowel syndrome, pancreatitis, small bowel obstruction,
 CC gastroesophageal reflux, duodenogastric reflux, Cushing's syndrome,
 CC hyperparathyroidism, Graves' disease, diabetic neuropathy, Paget's
 CC disease, polycystic ovary disease, orthostatic hypotension, postprandial
 CC hypotension, panic attacks, diabetes mellitus, hyperlipidaemia, insulin
 CC insensitivity, syndrome X, peptic ulcers, arthritis, obesity, opioid
 CC overdose and gastrointestinal bleeding.
 XX
 SQ Sequence 8 AA;

Query Match 53.0%; Score 35; DB 22; Length 8;
 Best Local Similarity 57.1%; Pred. No. 6.4e+05;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 cywgcg 7
 I:| | | |
 Db 2 cfwkcty 8

RESULT 11
 AAY51899
 ID AAY51899 standard; peptide; 6 AA.
 AC AAY51899;
 XX
 DT 16-JUN-2000 (first entry)
 XX
 DE Biostatin TT232 synthesising peptide 4.
 XX
 KW Solid phase synthesis; biostatin; cytostatic; tyrosine kinase inhibitor;
 KW cellular proliferation inhibition; somatostatin; antitumor.
 XX
 OS Unidentified.

Key Location/Qualifiers
 FT Modified-site 1 /note= "Ddz-Cys(Acm)"
 FT Modified-site 2 /note= "Tyr-(tBu)"
 FT Modified-site 3 /note= "D-form residue"
 FT Modified-site 4 /note= "Lys(2)"
 FT Modified-site 5 /note= "Cys(Acm)"
 FT Modified-site 6 /note= "Thr(tBu) with C-terminal amide group"
 XX
 PN WO200011032-A2.
 XX
 PD 02-MAR-2000.
 XX
 PF 20-AUG-1999; 99WO-EP06131.
 XX
 PR 20-AUG-1998; 98WO-EP05306.
 XX
 PA (ORPE-) ORPEGEN PHARMA GES BIOTECHNOLOGISCHE FOR.
 XX
 PI Braum G, Lifferth A, Birr C;
 XX
 DR WPI; 2000-224663/19.
 XX
 PT Biostatin preparation in high yield by solid synthesis, including
 PT closure of disulfide bridge before cleavage from support, useful as
 PT antitumor agent -
 XX
 PS Example 4; Page 25; 33pp; German.
 XX
 CC This invention describes a novel method for the solid phase (SP)
 CC synthesis of biostatin (TT 232) (I) which includes closing the disulfide
 CC bridge by oxidation of the completely or partially constructed peptide

CC while still bonded to the solid phase. The products of the invention
 CC have cytostatic activity and are tyrosine kinase inhibitors. (I) inhibits
 CC the tyrosine kinase activity of various human stomach cancer cell lines
 CC and thus inhibits cellular proliferation. The heptapeptide (I) described
 CC in the invention is a somatostatin analog which shows strong antitumor
 CC activity in vitro and in vivo. The SP synthesis method gives (I) is more
 CC easily and in markedly higher yield than by the method of EP0505680, in
 CC which the cyclization is carried out after cleaving the peptide from the
 CC resin. The solution method is also a simple synthesis of (I) in high
 CC yield; typically the tert-butyl-protected precursor can be oxidized in
 CC a yield of 70-80%. AAY51896-Y51900 represent peptides used in the
 CC synthesis of biostatin TT232 described in the method of the invention.
 XX
 SQ Sequence 6 AA;

Query Match 51.5%; Score 34; DB 21; Length 6;
 Best Local Similarity 80.0%; Pred. No. 6.4e+05;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 cywgc 5
 | | | |
 Db 1 cywkc 5

RESULT 12
 AAR27179
 ID AAR27179 standard; peptide; 7 AA.
 AC AAR27179;
 XX
 DT 21-MAY-1998 (first entry)
 XX
 DE Somatostatin analogue tyrosine kinase inhibitor #3.
 XX
 KW tumour; cancer; neoplasm; malignancy; psoriasis; regulation of;
 KW release of; growth hormone; insulin; glucagon; prolactin;
 KW inhibit exopeptidase; inhibit tumour growth; tumour transformation.
 XX
 OS Synthetic.

Key Location/Qualifiers
 FT misc_difference 1 /note= "D-form"
 FT disulfide_bond 2..6
 FT misc_difference 4
 FT modified_site 7 /note= "D-form"
 FT /note= "amidated"
 XX
 PN EP505680-A.
 XX
 PD 30-SEP-1992.
 XX
 PF 27-JAN-1992; 92EP-0101196.
 XX
 PR 25-JAN-1991; 91HU-0000272.
 XX
 PA (KERT/) KERT G.
 PA (BIOS-) BIOSIGNAL KUTATO FEJLESZTO KFT.
 XX
 PI Balogh A, Bokonyi G, Csuka O, Horvath A, Horvath J, Idel M;
 PI Keri G, Mezo I, Seprodi J, Szoke B, Teplan I, Vadasz Z;
 PI Boekoenyi G, Seproedi J, Szoke B, Mezo I, Teplan I;
 XX
 DR WPI; 1992-325051/40.
 XX
 PT New octa-peptide or hepta-peptide somatostatin analogues - as
 PT tyrosine kinase inhibitors for treating tumours and psoriasis,
 PT and for regulating hormone release
 XX
 PS Claim 5; Page 30; 36pp; English.
 XX

CC This peptide is a somatostatin analogue showing more advantageous
CC and selective pharmacological action than somatostatin. It has a
CC structure inhibiting the activity of exopeptidases. It can be used
CC in medicaments to inhibit tumour growth or the activity of tyrosine
CC kinase enzymes involved in tumour transformation. It is also useful
CC for regulating the release of growth hormone, insulin, glucagon and
CC prolactin. It may also be used to inhibit pathological processes
CC such as psoriasis, elicited by the pathological proliferation of
CC skin cells.

XX Sequence 7 AA;

Query Match 51.5%; Score 34; DB 13; Length 7;

Best Local Similarity 80.0%; Pred. NO. 6.4e+05;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 cywgc 5

||| |

Db 2 cywkc 6

RESULT 13

AAK56817
ID AAK56817 standard; peptide; 7 AA.

XX AC AAK56817;

XX 20-APR-1995 (first entry)

XX Somatostatin analog 41 - acts as agonist to inhibit tumour growth.

DE somatostatin; agonist; analog; accelerating tumour growth; trauma;
KW surgery; treatment; solid primary tumour; metastatic tumour;
KW inhibition; tissue trauma; ulcer.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "H-Phe; D-form residue"

FT Misc-difference 4 /note= "D-form residue"

FT Modified-site 8 /note= "amidated residue"

FT WO9418231-A.

XX 18-AUG-1994.

XX 08-FEB-1994; 94WO-0501412.

XX 11-FEB-1993; 93US-0016720.

XX (BIOM-) BIOMEASURE INC.

XX Bodgen AE, Moreau J;

XX WPI; 1994-279685/34.

XX Inhibiting accelerating tumour growth after trauma, esp surgery -
PT by treatment with somatostatin or its agonists, esp applied to
PT the site of trauma

XX Disclosure; Page 12; 33pp; English.

XX Somatostatin agonists include the somatostatin analogs shown in
CC AAR56777-848 and AAR60374-83. These peptides are used in the invention
CC to inhibit the accelerated growth of tumours. The growth is induced
CC surgically and the tumour is of epithelial origin, eg. lung or colon
CC tumours or partic. prostate or breast tumours, or melanoma. The
CC peptides are known inhibitors of cancer cell growth and are partic.
CC applied to the site of trauma, topically or subcutaneously.

XX SQ Sequence 7 AA;

Query Match 51.5%; Score 34; DB 15; Length 7;
Best Local Similarity 80.0%; Pred. NO. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 cywgc 5

||| |

Db 2 cywkc 6

RESULT 14

AAW18456

ID AAW18456 standard; peptide; 7 AA.

XX AC AAW18456;

XX 15-JUL-1997 (first entry)

XX Somatostatin analogue heptapeptide.

XX Growth hormone; tumour; Antarelix; Lanreotide; endocrine;
KW acromegaly.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "D-form residue"

FT Disulfide-bond 2..6

FT Misc-difference 4 /note= "D-form residue"

FT Modified-site 7 /label= OTHER

FT /note= "L-2-methyltryptophan amide"

XX WO9705167-A1.

XX 13-FEB-1997.

XX 17-JUL-1996; 96WO-EP03149.

XX 28-JUL-1995; 95IT-0M1670.

XX (DEGH/) DEGHENGI R.

XX Deghenghi R;

XX WPI; 1997-145622/13.

XX New somatostatin analogue peptide(s) - which inhibit the release of
PT growth hormone, used e.g. for the treatment and diagnosis of tumours
PT Claim 2; Page 15; 21pp; English.

XX The present sequence is claimed as a specific example of the
CC generic peptide formula A-Cys-B-D-Trp-Lys-C-Cys-R, in which the
CC two Cys residues are joined by a disulphide bond, A = D-2-alkyltrp,
CC D-betanaphthylalanine (D-betaNal) or D-Phe; B = Tyr or Phe; C = Val
CC or Thr (sic); and R = Thr-NH2 or L-alkyltrp-NH2; with the proviso that
CC when A is D-betaNal, B is not Tyr, C is not Val and R is not Thr-NH2.
CC This peptide is a somatostatin analogue which has activity in
CC inhibiting the release of growth hormone. It can be used for
CC the treatment of diseases characterised by an imbalance of growth
CC hormone. In particular, it can be used for the treatment of endocrine
CC tumours and acromegaly. The peptide can also be used as a diagnostic
CC agent for tumours. It is particularly active compared with known
CC somatostatin analogues (eg. Antarelix or Lanreotide) or somatostatin.
XX Sequence 7 AA;

Query Match 51.5%; Score 34; DB 18; Length 7;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 cywgc 5
Db 2 cywkc 6

Db 2 cywkc 6

Search completed: August 12, 2002, 17:14:32
Job time: 245 sec

RESULT 15

AAAY22068
ID AAAY22068 standard; peptide; 7 AA.

XX

AC AAAY22068;

XX

DT 27-AUG-1999 (first entry)

XX

DE Somatostatin agonist peptide.

XX

KW Somatostatin; agonist; inhibitor; fibrosis; kidney; lung; liver; skin;
chemotherapy; glomerulonephritis; diabetes; cirrhosis; nasal polyposis;
allograft rejection; infection; human immune deficiency virus; therapy;
central nervous system; intraocular fibrosis.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Modified-site 1

FT /note= "H-D-Phe"

FT Misc-difference 4

FT /note= "D-form residue"

FT Modified-site 8

FT /note= "amidated"

XX

PN W09808528-A1.

XX

PD 05-MAR-1998.

XX

PF 28-MAY-1997; 97WO-US08999.

XX

PR 30-AUG-1996; 96US-0705790.

XX

FA (BIOM-) BIOMEASURE INC.

XX

PI Culler MD, Kasprzyk PG;

XX

DR WPI; 1998-207029/18.

XX

PT Inhibiting fibrosis by administration of somatostatin or its
agonists - particularly in lung, liver, kidney and skin, or where
caused by chemotherapy

PT

XX

PS Disclosure; Page 10; 24pp; English.

XX

CC This sequence represents a somatostatin agonist, that can be used in the
method of the invention. The method is for the inhibition of fibrosis by
the administration of somatostatin or particularly its agonists. The
method is specifically used against fibrosis in the kidney, lung, liver
or skin (particularly where caused by chemotherapy but also the result of
glomerulonephritis, diabetes, cirrhosis, allograft rejection or infection
with human immune deficiency virus), also in the central nervous system
(e.g. intraocular fibrosis) or nose (e.g. nasal polyposis). The method
can be used therapeutically or prophylactically.

Sequence 7 AA;

Query Match 51.5%; Score 34; DB 19; Length 7;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 cywgc 5

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:14:59 ; Search time 20.43 Seconds
(without alignments)
9.565 Million cell updates/sec

Title: 10-071247-2
Perfect score: 66
Sequence: 1 cywgcgyw 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 48605

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCITUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	28	42.4	7	1	US-08-487-006-128
2	28	42.4	7	2	US-08-488-659A-128
3	28	42.4	8	4	US-09-082-279B-919
4	28	42.4	8	4	US-09-315-304B-919
5	27	40.9	6	2	US-08-282-980B-8
6	27	40.9	6	3	US-08-931-095-5
7	27	40.9	8	1	US-08-286-748B-7
8	27	40.9	8	4	US-09-484-318-1
9	27	40.9	8	4	US-09-484-318-8
10	27	40.9	8	4	US-09-484-319-1
11	27	40.9	8	4	US-09-484-319-8
12	27	40.9	8	4	US-09-484-320-1
13	27	40.9	8	4	US-09-484-320-8
14	27	40.9	8	4	US-09-484-321-1
15	27	40.9	8	4	US-09-484-321-8
16	27	40.9	8	4	US-09-484-323-1
17	27	40.9	8	4	US-09-484-323-8
18	27	40.9	8	4	US-09-325-769-1
19	27	40.9	8	4	US-09-325-769-2
20	27	40.9	8	4	US-08-586-670A-13
21	27	40.9	8	4	US-08-586-670A-14
22	27	40.9	8	4	US-09-636-170-1
23	27	40.9	8	4	US-09-636-170-8
24	27	40.9	8	4	US-09-637-518-1
25	27	40.9	8	4	US-09-637-518-8
26	26	39.4	5	1	US-07-946-237-4
27	26	39.4	5	2	US-08-530-566-10

28 26 39.4 5 4 US-09-195-726-10 Sequence 10, Appl
29 26 39.4 5 4 US-09-067-753-10 Sequence 10, Appl
30 26 39.4 6 1 US-08-487-006-139 Sequence 139, App
31 26 39.4 6 2 US-08-488-659A-139 Sequence 139, App
32 26 39.4 6 4 US-09-329-350-5 Sequence 5, Appl
33 26 39.4 7 4 US-09-329-350-8 Sequence 8, Appl
34 26 39.4 7 6 531889-61 Patent No. 531889
35 26 39.4 8 2 US-08-520-535-12 Sequence 12, Appl
36 26 39.4 8 2 US-08-811-492-110 Sequence 110, App
37 26 39.4 8 2 US-09-079-432-12 Sequence 12, Appl
38 25 37.9 6 1 US-08-483-434A-11 Sequence 11, Appl
39 25 37.9 6 4 US-08-476-134A-21 Sequence 21, Appl
40 25 37.9 6 6 519020-17 Patent No. 519020
41 25 37.9 6 6 5506208-18 Patent No. 5506208
42 25 37.9 8 1 US-08-249-371-7 Sequence 7, Appl
43 25 37.9 8 1 US-08-408-120-6 Sequence 6, Appl
44 25 37.9 8 1 US-08-408-120-11 Sequence 11, Appl
45 25 37.9 8 3 US-08-482-304-29 Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-08-487-006-128
; Sequence 128, Application US/08487006
; Patent No. 5641861
; GENERAL INFORMATION:
; APPLICANT: Dooley, Colette T.
; APPLICANT: Houghten, Richard A.
; TITLE OF INVENTION: No. 5641861el Mu Opioid Receptor Ligands:
; TITLE OF INVENTION: Agonists and Antagonists
; NUMBER OF SEQUENCES: 222
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,006
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-TP 1706
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 128:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..7
; OTHER INFORMATION: /note= "with the exception of Gly
; OTHER INFORMATION: In position 6, all amino acids are D-amino acids."
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 7
; OTHER INFORMATION: /note= "Amino acid is amidated at
; OTHER INFORMATION: the C-terminal."
US-08-487-006-128

Query Match 42.4%; Score 28; DB 1; Length 7;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ywgc 5
:||||
Db 4 WWC 7

RESULT 2
US-08-488-659A-128
Sequence 128, Application US/08488659A
Patent No. 5919897
GENERAL INFORMATION:
APPLICANT: Dooley, Colette T.
APPLICANT: Houghten, Richard A.
TITLE OF INVENTION: Mu Opioid Receptor Ligands:
TITLE OF INVENTION: Agonists and Antagonists
NUMBER OF SEQUENCES: 223
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,659A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-Tp 1705
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 128:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: linear
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..7
OTHER INFORMATION: /note= "With the exception of Gly
OTHER INFORMATION: in position 6, all amino acids are D-amino acids."
FEATURE:
NAME/KEY: Peptide
LOCATION: 7
OTHER INFORMATION: /note= "Amino acid is amidated at
OTHER INFORMATION: the C-terminal."
US-08-488-659A-128

Query Match 42.4%; Score 28; DB 2; Length 7;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ywgc 5
:||||
Db 4 WWC 7

RESULT 3
US-09-082-279B-919
Sequence 919, Application US/09082279B

Patent No. 6258782
GENERAL INFORMATION:
APPLICANT: Barney, Shawn
APPLICANT: Guthrie, Kelly
APPLICANT: Merutka, Gene
APPLICANT: Anwer, Mohamed
APPLICANT: Lambert, Dennis
TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
TITLE OF INVENTION: PHARMACOKINETIC PROPERTIES
FILE REFERENCE: 7872-043
CURRENT APPLICATION NUMBER: US/09/082,279B
CURRENT FILING DATE: 1998-05-20
NUMBER OF SEQ ID NOS: 1515
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 919
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Core polypeptide
US-09-082-279B-919

Query Match 42.4%; Score 28; DB 4; Length 8;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 wgcgy 7
:||||
Db 1 WWCY 5

RESULT 4
US-09-315-304B-919
Sequence 919, Application US/09315304B
Patent No. 6348568
GENERAL INFORMATION:
APPLICANT: Barney, S.
APPLICANT: Guthrie, K.
APPLICANT: Merutka, G.
APPLICANT: Anwer, M.
APPLICANT: Lambert, D.
TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
TITLE OF INVENTION: PROPERTIES
FILE REFERENCE: 7872-052
CURRENT APPLICATION NUMBER: US/09/315,304B
CURRENT FILING DATE: 1999-05-20
PRIOR APPLICATION NUMBER: 09/082,279
PRIOR FILING DATE: 1998-05-20
NUMBER OF SEQ ID NOS: 1667
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 919
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Core polypeptide
US-09-315-304B-919

Query Match 42.4%; Score 28; DB 4; Length 8;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 wgcgy 7
:||||
Db 1 WWCY 5

RESULT 5
US-08-282-980B-8
Sequence 8, Application US/08282980B
Patent No. 5932189

```
;
; GENERAL INFORMATION:
; APPLICANT: Dean, Richard T.
; APPLICANT: McBride, William
; APPLICANT: Lister-James, John
; TITLE OF INVENTION: Peptides
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,980B
; FILING DATE: 29-JUL-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5932189nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 92,385-I
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..6
; OTHER INFORMATION: /label= Cyclic
; OTHER INFORMATION: /note= "The amino terminus and carboxyl terminus
; OTHER INFORMATION: are linked by a covalent bond."
;
; QUERY MATCH 40.9%; Score 27; DB 2; Length 6;
; Best Local Similarity 100.0%; Pred. No. 1.7e+05;
; Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 CYW 3
; DB 1 CYW 3
;
; RESULT 6
; US-08-931-095-5
; Sequence 5, Application US/08931095
; Patent No. 6017512
; GENERAL INFORMATION:
; APPLICANT: Dean, Richard T.
; APPLICANT: McBride, William
; APPLICANT: Lister-James, John
; TITLE OF INVENTION: Radiolabeled Peptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive, 32nd Floor
; CITY: Chicago
```

```
;
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/931,095
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6017512nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 92,385-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..6
; OTHER INFORMATION: /label= Cyclic
; OTHER INFORMATION: /note= "The amino terminus and carboxyl terminus
; OTHER INFORMATION: are linked by a covalent bond; the amino terminus
; OTHER INFORMATION: is substituted with a methyl group."
;
; NAME/KEY: Modified-site
; LOCATION: 1..3
; OTHER INFORMATION: /label= Variant residues
; OTHER INFORMATION: /note= "The Trp is in the D conformation."
;
; US-08-931-095-5
;
; QUERY MATCH 40.9%; Score 27; DB 3; Length 6;
; Best Local Similarity 100.0%; Pred. No. 1.7e+05;
; Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 CYW 3
; DB 1 CYW 3
;
; RESULT 7
; US-08-286-748B-7
; Sequence 7, Application US/08286748B
; Patent No. 5759542
; GENERAL INFORMATION:
; APPLICANT: Victor Gurewich
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
; TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/286,748B
FILING DATE: August 5, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: J. Peter Fasse
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04547/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 8
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
OTHER INFORMATION: Xaa is D- -Naphthylalanine
US-08-286-748B-7

Query Match 40.9%; Score 27; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cyw 3
|||
Db 2 cyw 4

RESULT 8

US-09-484-318-1
; Sequence 1, Application US/09484318
; Patent No. 6180085
; GENERAL INFORMATION:
; APPLICANT: Achilefu, Samuel
; APPLICANT: Rajagopalan, Raghavan
; APPLICANT: Dorshow, Richard B
; APPLICANT: Bugaj, Joseph E.
; TITLE OF INVENTION: NOVEL DYES
; FILE REFERENCE: DNA STRING
; CURRENT APPLICATION NUMBER: US/09/484,318
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Octreotide
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: This is D-phenylalanine
; NAME/KEY: DISULFID
; LOCATION: (2)..(7)
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: This is D-tryptophan
; NAME/KEY: SITE
; LOCATION: (8)
; OTHER INFORMATION: This C-terminal residue ends with a hydroxyl
US-09-484-318-1

Query Match 40.9%; Score 27; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cyw 3
|||
Db 2 cyw 4

RESULT 9

US-09-484-318-8
; Sequence 8, Application US/09484318
; Patent No. 6180085
; GENERAL INFORMATION:
; APPLICANT: Achilefu, Samuel
; APPLICANT: Rajagopalan, Raghavan
; APPLICANT: Dorshow, Richard B
; APPLICANT: Bugaj, Joseph E.
; TITLE OF INVENTION: NOVEL DYES
; FILE REFERENCE: DNA STRING
; CURRENT APPLICATION NUMBER: US/09/484,318
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Octreotate
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: This is D-phenylalanine
; NAME/KEY: DISULFID
; LOCATION: (2)..(7)
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: This is D-tryptophan
US-09-484-318-8

Query Match 40.9%; Score 27; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cyw 3
|||
Db 2 cyw 4

RESULT 10

US-09-484-319-1
; Sequence 1, Application US/09484319
; Patent No. 6180086
; GENERAL INFORMATION:
; APPLICANT: Achilefu, Samuel
; APPLICANT: Rajagopalan, Raghavan
; APPLICANT: Dorshow, Richard B
; APPLICANT: Bugaj, Joseph E.
; TITLE OF INVENTION: NEW HYDROPHILIC CYANINE DYES
; FILE REFERENCE: DNA STRING
; CURRENT APPLICATION NUMBER: US/09/484,319
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Octreotide
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: This is D-phenylalanine
; NAME/KEY: DISULFID
; LOCATION: (2)..(7)
; NAME/KEY: SITE


```
; LOCATION: (4)
; OTHER INFORMATION: This is D-tryptophan
; NAME/KEY: SITE
; LOCATION: (8)
; OTHER INFORMATION: This C-terminal residue ends with a hydroxyl
US-09-484-319-1

Query Match          40.9%; Score 27; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cyw 3
   |||
Db 2 CYW 4

RESULT 11
US-09-484-319-8
; Sequence 8, Application US/09484319
; Patent No. 6180086
; GENERAL INFORMATION:
; APPLICANT: Achillefu, Samuel
; APPLICANT: Rajagopalan, Raghavan
; APPLICANT: Dorshow, Richard B
; APPLICANT: Bugaj, Joseph E.
; TITLE OF INVENTION: NEW HYDROPHILIC CYANINE DYES
; FILE REFERENCE: DNA STRING
; CURRENT APPLICATION NUMBER: US/09/484,319
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Octreotate
; LOCATION: (1)
; NAME/KEY: SITE
; OTHER INFORMATION: This is D-phenylalanine
; NAME/KEY: DISULFID
; LOCATION: (2)..(7)
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: This is D-tryptophan
US-09-484-319-8

Query Match          40.9%; Score 27; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cyw 3
   |||
Db 2 CYW 4

RESULT 12
US-09-484-320-1
; Sequence 1, Application US/09484320
; Patent No. 6180087
; GENERAL INFORMATION:
; APPLICANT: Achillefu, Samuel
; APPLICANT: Rajagopalan, Raghavan
; APPLICANT: Dorshow, Richard B
; APPLICANT: Bugaj, Joseph E.
; TITLE OF INVENTION: Tunable Indocyanine Dyes For Biomedical Applications
; FILE REFERENCE: dnastng
; CURRENT APPLICATION NUMBER: US/09/484,320
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Octreotate
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: This is D-phenylalanine
; NAME/KEY: DISULFID
; LOCATION: (2)..(7)
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: This is D-tryptophan
US-09-484-319-8

Query Match          40.9%; Score 27; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cyw 3
   |||
Db 2 CYW 4

RESULT 13
US-09-484-320-8
; Sequence 8, Application US/09484320
; Patent No. 6180087
; GENERAL INFORMATION:
; APPLICANT: Achillefu, Samuel
; APPLICANT: Rajagopalan, Raghavan
; APPLICANT: Dorshow, Richard B
; APPLICANT: Bugaj, Joseph E.
; TITLE OF INVENTION: Tunable Indocyanine Dyes For Biomedical Applications
; FILE REFERENCE: dnastng
; CURRENT APPLICATION NUMBER: US/09/484,320
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Octreotate
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: This is D-phenylalanine
; NAME/KEY: DISULFID
; LOCATION: (2)..(7)
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: This is D-tryptophan
US-09-484-320-8

Query Match          40.9%; Score 27; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cyw 3
   |||
Db 2 CYW 4

RESULT 14
US-09-484-321-1
; Sequence 1, Application US/09484321
```

```

; Patent No. 6183726
; GENERAL INFORMATION:
; APPLICANT: Achilefu, Samuel
; APPLICANT: Rajagopalan, Raghavan
; APPLICANT: Dorshow, Richard B
; APPLICANT: Bugaj, Joseph E.
; TITLE OF INVENTION: VERSATILE HYDROPHILIC DYES
; FILE REFERENCE: DNA STRING
; CURRENT APPLICATION NUMBER: US/09/484,321
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Octreotide
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: This is D-phenylalanine
; NAME/KEY: DISULFID
; LOCATION: (2)..(7)
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: This is D-tryptophan
; NAME/KEY: SITE
; LOCATION: (8)
; OTHER INFORMATION: This C-terminal residue ends with a hydroxyl
US-09-484-321-1

Query Match      40.9%; Score 27; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 cyw 3
Db      2 cyw 4

```

```

; Patent No. 6183726
; GENERAL INFORMATION:
; APPLICANT: Achilefu, Samuel
; APPLICANT: Rajagopalan, Raghavan
; APPLICANT: Dorshow, Richard B
; APPLICANT: Bugaj, Joseph E.
; TITLE OF INVENTION: VERSATILE HYDROPHILIC DYES
; FILE REFERENCE: DNA STRING
; CURRENT APPLICATION NUMBER: US/09/484,321
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Octreotate
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: This is D-phenylalanine
; NAME/KEY: DISULFID
; LOCATION: (2)..(7)
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: This is D-tryptophan
US-09-484-321-8

RESULT 15
US-09-484-321-8
; Sequence 8, Application US/09484321
; Patent No. 6183726
; GENERAL INFORMATION:
; APPLICANT: Achilefu, Samuel
; APPLICANT: Rajagopalan, Raghavan
; APPLICANT: Dorshow, Richard B
; APPLICANT: Bugaj, Joseph E.
; TITLE OF INVENTION: VERSATILE HYDROPHILIC DYES
; FILE REFERENCE: DNA STRING
; CURRENT APPLICATION NUMBER: US/09/484,321
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Octreotate
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: This is D-phenylalanine
; NAME/KEY: DISULFID
; LOCATION: (2)..(7)
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: This is D-tryptophan
US-09-484-321-1

```

```

Query Match      40.9%; Score 27; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 cyw 3
Db      2 cyw 4

Search completed: August 12, 2002, 17:15:00
Job time: 163 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2002, 17:18:23 ; Search time 51.76 Seconds
(without alignments)
17.168 Million cell updates/sec

Title: 10-071247-1
Perfect score: 59
Sequence: 1 gywgkgyw 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq_032802.*
1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	ID	Description
1	59	100.0	8	AA1903714	Fluorine-18 (F-18)
2	59	100.0	8	AA1903716	Fluorine-18 (F-18)
3	59	100.0	9	AA1903716	Immunogenic peptid
4	51	86.4	8	AA1903716	Immunogenic peptid
5	48	81.4	8	AA1903716	Immunogenic peptid
6	47	79.7	1365	AA1903716	H2 homologue of pr
7	47	79.7	1365	AA1903716	KRE5. Saccharomyc
8	45	76.3	8	AA1903715	Fluorine-18 (F-18)
9	45	76.3	8	AA1903715	Immunogenic peptid
10	45	76.3	164	AA1903715	Human novel secret
11	44	74.6	152	AA1903715	Human protein seq

12	44	74.6	1090	22	AA1903714	Human protein seq
13	44	74.6	1180	22	AA1903714	Novel human diagno
14	42	71.2	121	22	AA1903714	Anti-adipocyte mon
15	42	71.2	126	22	AA1903714	Drosophila melanog
16	42	71.2	1329	21	AA1903714	Streptococcus pneu
17	41	69.5	844	19	AA1903714	Human ligase IV.
18	41	69.5	911	17	AA1903714	DNA-ligase-III. H
19	41	69.5	1447	22	AA1903714	UGT. Schizosacch
20	41	69.5	1891	22	AA1903714	Novel human diagno
21	40	67.8	139	21	AA1903714	Human 5' EST relat
22	40	67.8	142	18	AA1903714	Pe1B-VH4715 fusion
23	40	67.8	274	18	AA1903714	VH4715-linker-VH34
24	40	67.8	282	18	AA1903714	HindIII-EcoRI inse
25	40	67.8	282	18	AA1903714	VH4715-linker-VL47
26	40	67.8	289	18	AA1903714	HindIII-EcoRI inse
27	40	67.8	358	10	AA1903714	Muscarinic M3 rece
28	39	66.1	532	18	AA1903714	HindIII-EcoRI inse
29	39	66.1	45	22	AA1903714	Human secreted pro
30	39	66.1	52	21	AA1903714	Human secreted pro
31	39	66.1	261	14	AA1903714	Chimeric Ig superf
32	39	66.1	267	14	AA1903714	Chimeric Ig superf
33	39	66.1	289	21	AA1903714	Arabidopsis thalia
34	39	66.1	340	21	AA1903714	Arabidopsis thalia
35	39	66.1	396	21	AA1903714	Arabidopsis thalia
36	39	66.1	655	13	AA1903714	Sequence encoded b
37	39	66.1	700	21	AA1903714	Yeast G protein-co
38	39	66.1	715	13	AA1903714	Sequence encoded b
39	39	66.1	847	21	AA1903714	Yeast G protein-co
40	39	66.1	864	21	AA1903714	Yeast G protein-co
41	39	66.1	953	21	AA1903714	Yeast G protein-co
42	39	66.1	953	21	AA1903714	Yeast G protein-co
43	39	66.1	961	21	AA1903714	Yeast G protein-co
44	39	66.1	1197	22	AA1903714	Novel human diagno
45	39	66.1	1197	22	AA1903714	Novel human diagno

ALIGNMENTS

RESULT 1
AA1903714
ID AAY03714 standard; peptide; 8 AA.
XX AC AAY03714;
XX AC AAY03714;
DT 08-JUN-1999 (first entry)
XX DE Fluorine-18 (F-18) labeled peptide 1.
XX DE 18F radionuclide; targeting vector; positron emission tomography; F-18;
KW KW radiolabeling; thiol; fluorine-18.
XX OS Synthetic.
XX OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 1 /note= "optionally has a free or protected
FT amino acid group"
FT FT
FT Misc-difference 2 /note= "D-form residue"
FT FT
FT Misc-difference 3 /note= "D-form residue"
FT FT
FT Misc-difference 5 /note= "D-form residue"
FT FT
FT Misc-difference 7 /note= "D-form residue; optionally has a free or
FT Protected amino acid group"
FT Misc-difference 8 /note= "D-form residue"
FT FT
FT Misc-difference 8 /note= "D-form residue"
FT FT
PN WO9911590-A1.
XX 11-MAR-1999.
PD 11-MAR-1999.

```
XX 03-SEP-1998; 98WO-US18268.
PF
XX
PR 03-SEP-1997; 97US-0057485.
XX
PA (IMMU-) IMMUNOMEDICS INC.
XX
PI Griffiths GL;
XX
XX WPI; 1999-228967/19.
XX
XX Radiolabeling thiol-containing peptides with fluorine-18
XX
XX Claim 13; Page 15; 22pp; English.
XX
XX The invention relates to a method for incorporating 18F radionuclide
XX into peptide-containing targeting vectors for use in clinical positron
XX emission tomography. Radiolabeling thiol-containing peptides with
XX fluorine-18 (F-18) comprises reacting a peptide comprising a free thiol
XX group with a labeling reagent of formula: 18F-(CH2)m-CR1R2-(CH2)n-X, or
XX a fluorinated alkene in which at least one of the two double bonded
XX carbon atoms bears at least one leaving group comprising I, Br, Cl,
XX azide, tosylate, mesylate, nosylate or triflate. n, m = 0-2; n+m = 0-2;
XX x = I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, maleimide
XX (optionally substituted by 1-2 alkyl) or 3-sulfomaleimide; R1, R2 = I,
XX Br, Cl, azide, tosylate, mesylate, nosylate, triflate, H, CONH2, COOH,
XX OH, sulfonic acid, tertiary amine, quaternary ammonium, alkyl
XX (optionally substituted by CONH2, COOH, OH, sulfonic acid, tertiary amine
XX or quaternary ammonium), COOR', CONR'2 or COR'; and R' = 1-6C alkyl or
XX phenyl. The method is used for radiolabeling peptide-containing
XX targeting vectors such as proteins, antibodies, antibody fragments and
XX receptor-targeted peptides for use in routine clinical positron
XX emission tomography. The method is simple and efficient. The method uses
XX the unique property of the free thiol groups which are rapidly alkylated
XX at neutral pH and moderate temperature. Sequences AAY03714-716 represent
XX examples of F-18 labeled peptides used in the method of detecting a
XX tissue.
XX
XX Sequence 8 AA;
SQ
Query Match 100.0%; Score 59; DB 20; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gywgkgyw 8
Db 1 gywgkgyw 8
|||||
|||||

RESULT 2
AAY03716 100.0%; Score 59; DB 20; Length 8;
ID AAY03716 standard; peptide; 8 AA.
XX
XX AAY03716;
XX
XX 08-JUN-1999 (first entry)
XX
XX Fluorine-18 (F-18) labeled peptide 3.
XX
XX 18F radionuclide; targeting vector; positron emission tomography; F-18;
KW radiolabeling; thiol; Fluorine-18.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Modified-site 1
XX /note= "N-terminal acetylation"
XX
XX Misc-difference 2
XX /note= "D-form residue; D-Iodo-Tyr"
XX
XX Misc-difference 3
XX /note= "D-form residue"
XX
XX Misc-difference 5
```

```
FT Misc-difference 7 /note= "D-form residue; optionally acetylated"
FT FT
FT Misc-difference 8 /note= "D-form residue; D-Iodo-Tyr"
FT FT
FT Misc-difference 8 /note= "D-form residue"
XX
XX WO9911590-A1.
XX
XX 11-MAR-1999.
XX
XX 03-SEP-1998; 98WO-US18268.
XX
XX 03-SEP-1997; 97US-0057485.
XX
XX (IMMU-) IMMUNOMEDICS INC.
XX
XX Griffiths GL;
XX
XX WPI; 1999-228967/19.
XX
XX Radiolabeling thiol-containing peptides with fluorine-18
XX
XX Claim 15; Page 15; 22pp; English.
XX
XX The invention relates to a method for incorporating 18F radionuclide
XX into peptide-containing targeting vectors for use in clinical positron
XX emission tomography. Radiolabeling thiol-containing peptides with
XX fluorine-18 (F-18) comprises reacting a peptide comprising a free thiol
XX group with a labeling reagent of formula: 18F-(CH2)m-CR1R2-(CH2)n-X, or
XX a fluorinated alkene in which at least one of the two double bonded
XX carbon atoms bears at least one leaving group comprising I, Br, Cl,
XX azide, tosylate, mesylate, nosylate or triflate. n, m = 0-2; n+m = 0-2;
XX x = I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, maleimide
XX (optionally substituted by 1-2 alkyl) or 3-sulfomaleimide; R1, R2 = I,
XX Br, Cl, azide, tosylate, mesylate, nosylate, triflate, H, CONH2, COOH,
XX OH, sulfonic acid, tertiary amine, quaternary ammonium, alkyl
XX (optionally substituted by CONH2, COOH, OH, sulfonic acid, tertiary amine
XX or quaternary ammonium), COOR', CONR'2 or COR'; and R' = 1-6C alkyl or
XX phenyl. The method is used for radiolabeling peptide-containing
XX targeting vectors such as proteins, antibodies, antibody fragments and
XX receptor-targeted peptides for use in routine clinical positron
XX emission tomography. The method is simple and efficient. The method uses
XX the unique property of the free thiol groups which are rapidly alkylated
XX at neutral pH and moderate temperature. Sequences AAY03714-716 represent
XX examples of F-18 labeled peptides used in the method of detecting a
XX tissue.
XX
XX Sequence 8 AA;
SQ
Query Match 100.0%; Score 59; DB 20; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gywgkgyw 8
Db 1 gywgkgyw 8
|||||
|||||

RESULT 3
AAY76816
ID AAY76816 standard; peptide; 9 AA.
XX
XX AAY76816;
XX
XX 28-APR-2000 (first entry)
XX
XX Immunogenic peptide for bi-specific antibody recognition.
XX
XX Immunogenic peptide; bi-specific antibody; diagnosis; immune response;
XX diseased tissue identification; therapy.
XX
XX Synthetic.
OS
```

XX	Key	Location/Qualifiers	
FH	Misc-difference 1	/note= "free amino acid group, protected amino acid group, chelating agent or a metal-chelate complex"	
FT			
FT			
FT			
FT			
FT	Misc-difference 3	/note= "D-form residue"	
FT			
FT	Misc-difference 4	/note= "D-form residue"	
FT			
FT	Misc-difference 6	/note= "D-form residue; modified with free amino acid group, protected amino acid group, chelating agent or a metal-chelate complex"	
FT			
FT	Misc-difference 8	/note= "D-form residue"	
FT			
FT	Misc-difference 9	/note= "D-form residue"	
FT			
XX	W09966951-A2.		
PN			
XX			
XX	29-DEC-1999.		
XX			
XX	22-JUN-1999;	99WO-US13879.	
XX			
XX	22-JUN-1998;	98US-0090142.	
PR			
PR	14-OCT-1998;	98US-0104156.	
XX			
XX	(IMMU-) IMMUNOMEDICS INC.		
PA			
XX			
XX	Hansen HJ, Griffiths GL, Leung S, McBride WJ, Qu Z;		
PI			
XX	WPI; 2000-160561/14.		
XX			
DR			
XX			
PT	Bi-specific antibodies that bind specific target tissue and targeted conjugates		
PT			
XX			
XX	Claim 21; Page 61; 76pp; English.		
XX			
CC	This sequence represents an immunogenic peptide for a bi-specific antibody. The invention relates to a method of treating or identifying diseased tissues in a patient comprising administering a bi-specific antibody (or fragment) having at least 1 arm (A) that specifically binds a targeted tissue and at least 1 arm (B) that specifically binds a targetable conjugate. The methods and bi-specific antibodies and fusion proteins are useful for pre-targeting methods of diagnosis and therapy. It is advantageous to raise bi-specific antibodies against a targetable conjugate that is capable of carrying at least 1 diagnostic or therapeutic agent. The characteristics of the chelator, metal chelate complex, therapeutic agent or diagnostic agent can be varied to accommodate differing applications without raising new bi-specific antibodies for each new application. The targetable conjugate is selected to elicit sufficient immune responses and also for rapid in vivo clearance when used within the bi-specific antibody targeting method.		
CC			
XX	Sequence	9 AA;	
SQ			
Query Match 100.0%; Score 59; DB 21; Length 9;			
Best Local Similarity 100.0%; Pred. No. 6.4e+05;			
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 gywgkgyw 8		
Db	2 gywgkgyw 9		
RESULT 4			
AAV76818			
ID	AAV76818 standard; peptide; 8 AA.		
XX			
AC	AAV76818;		
XX			

ID AAB18511 standard; protein; 841 AA.
XX AC AAB18511;
XX DT 15-JAN-2001 (first entry)
XX DE H2 homologue of prolyl-tripeptidyl peptidase DPP.
XX KW Prolyl tripeptidyl-peptidase; amidolytic activity; periodontal disease;
XX KW gingivitis; periodontitis.
XX OS Porphyromonas gingivalis.
XX PN WO200052147-A2.
XX PD 08-SEP-2000.
XX PF 03-MAR-2000; 2000WO-US05551.
XX PR 05-MAR-1999; 99US-0123148.
XX PA (UYGE-) UNIV GEORGIA RES FOUND INC.
XX PA (TRAV/) TRAVIS J.
XX PA (POTE/) POTEPA J.
XX PA (BANB/) BANBULA A.
XX PI Travis J, Potempa J, Banbula A;
XX DR WPI; 2000-594181/56.
XX PT Prolyl tripeptidyl-peptidase, active analog, fragment or variant useful
XX PT for identifying its inhibitor which is useful for protecting an animal
XX PT from a periodontal disease such as gingivitis and periodontitis -
XX PS Claim 22; Fig 6; 58pp; English.
XX CC The present sequence represents a H2 homologue of a prolyl
CC tripeptidyl-peptidase (designated DPP) from Porphyromonas gingivalis.
CC CC The prolyl tripeptidyl-peptidase has an amidolytic activity, and cleaves
CC a peptide bond in a target polypeptide having at least 4 amino acids.
CC CC This bond is between a proline and an amino acid attached to the
CC alpha-carboxyl group end of the proline. The polypeptide is useful
CC for identifying inhibitors. These inhibitors are then useful for
CC reducing the growth of bacterium or for protecting an animal from a
CC periodontal disease such as gingivitis and periodontitis caused by
CC Porphyromonas gingivalis.
XX SQ Sequence 841 AA;

Query Match 81.4%; Score 48; DB 21; Length 841;
Best Local Similarity 85.7%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywgkgyw 8
||| |||
Db 719 ywsggyw 725

RESULT 6
AAR36780
ID AAR36780 standard; protein; 1365 AA.
XX AC AAR36780;
XX DT 16-JUL-1993 (first entry)
XX DE KRE5.
XX KW Yeast; cell wall; beta-glucan; assembly; pathway; KRE1; KRE5; growth;
XX KW secretory; O-linked mannose; (1-6)-beta-glucan; epistasis; morphology;
XX KW hydrophilic; glycoprotein; COOH-terminal; endoplasmic reticulum; ER;
XX KW retention signal; antifungal agent.

XX OS Saccharomyces cerevisiae.
XX PN US5194600-A.
XX PD 16-MAR-1993.
XX PF 05-MAR-1990; 90US-0488316.
XX PR 05-MAR-1990; 90US-0488316.
XX PA (ROYA-) ROYAL INST ADVANCEMENT LEARNING.
XX PI Boone C, Bussey H, Hill K, Meaden P, Sommer SS;
XX DR WPI; 1993-109384/13.
XX DR N-PSDB; AAQ38899.
XX PT New DNA encoding genes which participate in beta-glucan assembly
XX PT - useful for producing mutants for in-vivo screening of
XX PT antifungal agents and providing tools for in-vitro screening
XX PS Claim 1; Columns 38-44; 24pp; English.
XX CC The sequences given in AAR34785 and AAR36780 represent proteins which
CC participate in a yeast cell wall beta-glucan assembly pathway.
CC CC These proteins represent KRE1 and KRE5 respectively, and are
CC essential for normal cell growth. KRE1 is a Ser/Thr rich protein
CC that is directed into the yeast secretory pathway, where it is
CC highly modified, probably through addition of O-linked mannose
CC residues. Gene disruption of the KRE1 locus leads to a 40% reduced
CC level of cell wall (1-6)-beta-glucan. Mutations at KRE5 also caused
CC defects in cell wall (1-6)-beta-glucan production and appears to be
CC epistatic to KRE1. KRE5 is a large hydrophilic secretory glyco-
CC protein which contains the COOH-terminal endoplasmic reticulum (ER)
CC retention signal (His-Asp-Glu-Leu). Deletion of the KRE5 gene results
CC in cells with aberrant morphology and extremely compromised growth.
CC KRE1 and KRE5 are useful as tools for the in vitro screening of anti-
CC fungal agents which inhibit fungi pathogenic to plants and animals.
CC The genes can be used to produce mutants for in vivo screening of
CC antifungal agents.
XX SQ Sequence 1365 AA;

Query Match 79.7%; Score 47; DB 14; Length 1365;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 gywkggyw 8
||| :|||
Db 1234 gywkegyw 1241

RESULT 7
AAB72446
ID AAB72446 standard; protein; 1365 AA.
XX AC AAB72446;
XX DT 08-MAY-2001 (first entry)
XX DE Yeast KRE5.
XX KW UGGT; UDP-glucose:glycoprotein glucosyltransferase; enzyme; KRE5.
XX OS Saccharomyces cerevisiae.
XX PN WO200112845-A1.
XX PD 22-FEB-2001.
XX PF 27-JUL-2000; 2000WO-CA00883.

XX 18-AUG-1999; 99US-0376330.
 XX (CANA) NAT RES COUNCIL CANADA.
 XX
 PI Tessier DC, Dignard D, Bergeron JMM, Thomas DY;
 XX
 XX WPI; 2001-218358/22.
 XX
 PT Determining the effect of a test sample on UDP-glucose:glycoprotein
 PT glucosyltransferase (UGGT), useful for measuring UGGT activity,
 PT comprises exposing an acceptor substrate for UGGT to a labeled donor in
 PT the presence of UGGT -
 XX
 XX Disclosure; Fig 9; 95pp; English.
 XX
 CC The present invention relates to a method for determining the effect of a
 CC test sample on UDP-glucose:glycoprotein glucosyltransferase (UGGT)
 CC activity. The method comprises exposing an acceptor substrate for UGGT to
 CC a labelled donor in the presence of the test sample and UGGT. The method
 CC is useful for determining UGGT activity. In particular, the method is
 CC useful in glucosyltransferase assay and kinetics measurement for
 CC determining UGGT activity. UGGT is a soluble enzyme of the endoplasmic
 CC reticulum which catalyses the addition of a glucose residue onto
 CC asparagine-linked oligosaccharides, which are present on incorrectly
 CC folded glycoproteins. The present sequence was used in a sequence
 CC homology comparison with rat UGGT (see AAF60732 and AAB72436) which was
 CC used in the method of the present invention.
 XX
 XX Sequence 1365 AA;

Query Match 79.7%; Score 47; DB 22; Length 1365;
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 gywkggyw 8
 |||:||||
 Db 1234 gywkegyw 1241

RESULT 8
 AAY03715
 ID AAY03715 standard; peptide; 8 AA.
 XX
 AC AAY03715;
 XX
 DT 08-JUN-1999 (first entry)
 XX
 DE Fluorine-18 (F-18) labeled peptide 2.
 XX
 KW 18F radionuclide; targeting vector; positron emission tomography; F-18;
 KW radiolabeling; thiol; fluorine-18.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1 /note= "N-terminal acetylation; optionally has a free
 FT or protected thiol group"
 FT
 FT Misc-difference 2 /note= "D-form residue"
 FT
 FT Misc-difference 3 /note= "D-form residue"
 FT
 FT Misc-difference 5 /note= "D-form residue"
 FT
 FT Misc-difference 7 /note= "D-form residue; optionally has a free or
 FT protected thiol group"
 FT
 FT Misc-difference 8 /note= "D-form residue"
 FT
 FT Misc-difference 9 /note= "D-form residue"
 FT
 PN W09911590-A1.

XX 11-MAR-1999.
 PD
 XX
 XX 03-SEP-1998; 98WO-US18268.
 PF
 XX
 PR 03-SEP-1997; 97US-0057485.
 PR
 XX (IMMU-) IMMUNOMEDICS INC.
 PA
 XX Griffiths GL;
 PI
 XX WPI; 1999-228967/19.
 DR
 XX
 XX Radiolabeling thiol-containing peptides with fluorine-18
 PT
 XX
 PS Claim 14; Page 15; 22pp; English.

CC The invention relates to a method for incorporating 18F radionuclide
 CC into peptide-containing targeting vectors for use in clinical positron
 CC emission tomography. Radiolabeling thiol-containing peptides with
 CC fluorine-18 (F-18) comprises reacting a peptide comprising a free thiol
 CC group with a labeling reagent of formula: 18F-(CH₂)_m-CR1R2-(CH₂)_n-X, or
 CC a fluorinated alkene in which at least one of the two double bonded
 CC carbon atoms bears at least one leaving group comprising I, Br, Cl,
 CC azide, tosylate, mesylate, nosylate or triflate, n = 0-2; n+m = 0-2;
 CC X = I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, maleimide
 CC (optionally substituted by 1-2 alkyl) or 3-sulfonamideimide; R1, R2 = I,
 CC Br, Cl, azide, tosylate, mesylate, nosylate, triflate, H, CONH₂, COOH,
 CC OH, sulfonic acid, tertiary amine, quaternary ammonium, alkyl
 CC (optionally substituted by CONH₂, COOH, OH, sulfonic acid, tertiary amine
 CC or quaternary ammonium), COOR', CONR'2 or COR'; and R' = 1-6C alkyl or
 CC phenyl. The method is used for Radiolabeling peptide-containing
 CC targeting vectors such as proteins, antibodies, antibody fragments and
 CC receptor targeted peptides for use in routine clinical positron
 CC emission tomography. The method is simple and efficient. The method uses
 CC the unique property of the free thiol groups which are rapidly alkylated
 CC at neutral pH and moderate temperature. Sequences AAY03714-716 represent
 CC examples of F-18 labeled peptides used in the method of detecting a
 CC tissue.

XX Sequence 8 AA;

Query Match 76.3%; Score 45; DB 20; Length 8;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywgkggyw 8
 ||| |||
 Db 2 ywgcggyw 8

RESULT 9
 AAY76817
 ID AAY76817 standard; peptide; 8 AA.
 XX
 AC AAY76817;
 XX
 DT 28-APR-2000 (first entry)
 XX
 DE Immunogenic peptide for bi-specific antibody recognition.
 XX
 KW Immunogenic peptide; bi-specific antibody; diagnosis; immune response;
 KW diseased tissue identification; therapy.
 XX
 XX Synthetic.
 OS
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1 /note= "acetylated; modified with free amino acid group,
 FT protected amino acid group, chelating agent or a
 FT metal-chelate complex"
 FT
 FT Misc-difference 2

FT Misc-difference 3 /note= "D-form residue"
FT Misc-difference 5 /note= "D-form residue"
FT Misc-difference 5 /note= "D-form residue; modified with free amino acid
FT group, protected amino acid group, chelating
FT agent or a metal-chelate complex"
FT Misc-difference 7 /note= "D-form residue"
FT Misc-difference 8 /note= "D-form residue"
FT
XX WO9966951-A2.
XX 29-DEC-1999.
XX
XX 22-JUN-1999; 99WO-US13879.
XX
XX 22-JUN-1998; 98US-0090142.
XX 14-OCT-1998; 98US-0104156.
XX (IMMU-) IMMUNOMEDICS INC.
XX Hansen HU, Griffiths GL, Ieung S, McBride WJ, Qu Z;
PI WPI; 2000-160561/14.
XX
XX Bi-specific antibodies that bind specific target tissue and targeted
XX conjugates -
XX
XX Claim 22; Page 61; 76pp; English.
XX
XX This sequence represents an immunogenic peptide for a bi-specific
XX antibody. The invention relates to a method of treating or identifying
XX diseased tissues in a patient comprising administering a bi-specific
XX antibody (or fragment) having at least 1 arm (A) that specifically binds a
XX targeted tissue and at least 1 arm (B) that specifically binds a
XX targetable conjugate. The methods and bi-specific antibodies and fusion
XX proteins are useful for pre-targeting methods of diagnosis and therapy.
XX It is advantageous to raise bi-specific antibodies against a targetable
XX conjugate that is capable of carrying at least 1 diagnostic or
XX therapeutic agent. The characteristics of the chelator, metal chelate
XX complex, therapeutic agent or diagnostic agent can be varied to
XX accommodate differing applications without raising new bi-specific
XX antibodies for each new application. The targetable conjugate is selected
XX to elicit sufficient immune responses and also for rapid in vivo
XX clearance when used within the bi-specific antibody targeting method.
XX
XX Sequence 8 AA;
XX
XX Query Match 76.3%; Score 45; DB 21; Length 8;
XX Best Local Similarity 85.7%; Pred. No. 6.4e+05;
XX Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 ywgkgyw 8
Dd ||| |||
Dd 2 ywgcyw 8
RESULT 10
AAU16277
ID AAU16277 standard; Protein; 164 AA.
XX
XX AAU16277;
AC
XX
XX 07-NOV-2001 (first entry)
XX
XX Human novel secreted protein, Seq ID 1230.
XX
XX Human; immunosuppressive; antiarthritic; antirheumatic;
XX cytostatic; cardiatic; vasotropic; cerebroprotective; nootropic;
XX neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
XX
KW vulnerary; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin ageing; food additive; preservative; antiproliferative.
XX
OS Homo sapiens.
XX
XX WO200155322-A2.
XX
XX 02-AUG-2001.
PD
XX
PF 17-JAN-2001; 2001WO-US01341.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.

```
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246539.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.

PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251889.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-488783/53.
DR N-PSDB; AAS26264.
DR
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
XX Claim 11; SEQ ID No 1230; 980pp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC are also used in diagnosing a pathological condition or susceptibility
CC to a pathological condition. Antibodies to the proteins can also
CC be used in alleviating symptoms associated with the disorders and in
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC immunosorbant assays (ELISA). Disorders which are diagnosed or treated
CC include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC and ocular disorders e.g. corneal infection, and many other
CC disorders listed in the specification. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention.

Query Match 76.3%; Score 45; DB 22; Length 164;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 gywgkgyw 8
   :||| |
Db 125 ghwgkgew 132

RESULT 11
AAB99111
ID AAB99111 standard; Protein; 152 AA.
XX
AC AAB99111;
XX
DT 22-AUG-2001 (first entry)
XX
XX Human protein SEQ ID 4.
XX
XX Human; antiarthritic; cardiant; monoclonal antibody; keloid; arthritis;
KW Tumour Growth Factor-beta II receptor; tgf-beta II receptor; atopy;
KW signal transduction inhibition; tissue fibrosis; atherosclerosis.
XX
XX Homo sapiens.
XX
XX WO200136642-A1.
XX
XX 25-MAY-2001.
XX
```

PF 17-NOV-2000; 2000WO-JP08129.
XX
PR 18-NOV-1999; 99JP-0328681.
PR 08-NOV-2000; 2000JP-0340216.
XX
PA (NISB) JAPAN TOBACCO INC.
XX
PI Sakamoto S, Kamada M;
XX
XX WPI: 2001-343825/36.
DR N-PSDB; AAH41153.
XX
XX Human monoclonal antibodies recognizing human TGF-beta II receptor,
PT useful for treating TGF-beta associated diseases such as tissue
PT fibrosis -
XX
XX Claim 10; Page 94-95; 118pp; Japanese.
XX
CC The present invention relates to novel human monoclonal antibodies. The
CC antibodies can bind to human Tumour Growth Factor-beta (TGF-beta) II
CC receptor, resulting in the inhibition of the signal transduction of human
CC TGF-beta into cells. The antibodies can be used for the prevention and
CC treatment of diseases associated with the production of TGF-beta, such as
CC tissue fibrosis in the lung, liver, skin, kidney or other tissues, as
CC atherosclerosis, atopy, keloid and arthritis. The present sequence was
CC used in the present invention.
XX
SQ Sequence 152 AA;

Query Match 74.6%; Score 44; DB 22; Length 152;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gywkggyw 8
||| |
Db 118 gywgdgyw 125

RESULT 12
AAM80087
ID AAM80087 standard; Protein; 1090 AA.
XX
AC AAM80087;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3733.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.

c

PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
XX
XX Xue AJ, Yang Y, Wejhrman T, Goodrich R;
DR WPI: 2001-476283/51.
DR N-PSDB; AAK53220.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
PT
PS Claim 20; Page 420; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
XX were missing at the time of publication.
XX
SQ Sequence 1090 AA;

Query Match 74.6%; Score 44; DB 22; Length 1090;
Best Local Similarity 75.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gywkggyw 8
||| |
Db 1001 gywgdgyw 1008

RESULT 13
ABG19170
ID ABG19170 standard; Protein; 1180 AA.
XX
AC ABG19170;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #19161.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
DR WPI: 2001-639362/73.
DR N-PSDB; AAS83357.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT biodiversity -
 PS Claim 20; SEQ ID No 49529; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence . 1180 AA;

Query Match 74.6%; Score 44; DB 22; Length 1180;
 Best Local Similarity 75.0%; Pred. No. 2.7e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gywgkgyw 8
 |||||
 Db 1091 gywgdgew 1098

RESULT 14
 AAU02601
 ID AAU02601 standard; Protein; 121 AA.
 XX
 AC AAU02601;
 XX
 DT 29-AUG-2001 (first entry)
 XX
 DE Anti-adipocyte monoclonal antibody heavy chain, FAT 91.
 XX
 KW Antibody; adipocyte; heavy chain; light chain; obesity; fat;
 KW heart disease; complementarity determining region; CDR.
 XX
 OS Homo sapiens.
 XX
 PN WO200127279-A1.
 XX
 PD 19-APR-2001.
 XX
 PF 11-OCT-2000; 2000WO-GB03900.
 XX
 PR 12-OCT-1999; 99US-0158812.
 XX
 PA (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX
 PI Edwards BM, Main SH, Vaughan TJ;
 XX
 DR WPI; 2001-282031/29.
 DR N-PSDB; AAS03501.
 XX
 PT Panel of specific binding members of antibody molecules which bind to
 PT whole adipocytes is used in the treatment of obesity and obesity
 PT related diseases -
 XX
 PS Claim 1; Page 156; 182pp; English.

XX
 CC AAU02501-AAU02635, and AAU02641-AAU02748 represent the amino acid
 CC sequences of anti-adipocyte monoclonal antibody heavy chain, light
 CC chain, and heavy chain complementarity determining regions (CDR) of the
 CC invention. The antibodies can be used in the treatment of obesity and
 CC obesity related diseases. The antibodies can be used to deliver drugs or
 CC pro-drugs directly to the fat mass of an obese patient or the antibody
 CC can be used as a therapeutic itself. Antibodies binding specifically to
 CC adipocytes can be used to activate the immune system to destroy the cells
 CC by complement mediated lysis. The antibodies may be labeled with a
 CC detectable label such as radiolabel, fluorescent or chemical group and
 CC used in methods of diagnosis in human subjects e.g. to determine the
 CC presence of adipocyte antigen on the surface of an adipocyte to detect or
 CC determine the presence or level of adipocytes in a cell or tissue sample.
 CC The antibodies can be used as an alternative means of treatment for obese
 CC patients other than undergoing surgery to remove excess fat. Antibodies
 CC for different types of fat deposits can also be produced e.g. intra-
 CC abdominal fat associated with heart disease.
 XX
 SQ Sequence 121 AA;

Query Match 71.2%; Score 42; DB 22; Length 121;
 Best Local Similarity 71.4%; Pred. No. 59;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ywgkgyw 8
 |||||
 Db 110 ywgrgqw 116

RESULT 15
 ABB69676
 ID ABB69676 standard; Protein; 126 AA.
 XX
 AC ABB69676;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 35820.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL13779.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 35820; 21pp. + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
xx

SQ Sequence 126 AA;

Query Match 71.2%; Score 42; DB 22; Length 126;
Best Local Similarity 71.4%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 ywgkgyw 8

|||

Db 37 ywnsgyw 43

Search completed: August 12, 2002, 17:21:41
Job time: 198 sec